A mean–field statistical theory for the nonlinear Schrödinger equation

Richard Jordan *
T-7 and CNLS
MS B258
Los Alamos National Laboratory
Los Alamos, NM 87545

Bruce Turkington
Department of Mathematics and Statistics
University of Massachusetts
Amherst, MA 01003

Craig L. Zirbel
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, OH 43403

April 5, 1999

Abstract

A statistical model of self-organization in a generic class of one-dimensional nonlinear Schrödinger (NLS) equations on a bounded interval is developed. The main prediction of this model is that the statistically preferred state for such equations consists of a deterministic coherent structure coupled with fine-scale, random fluctuations, or radiation. The model is derived from equilibrium statistical mechanics by using a mean–field approximation of the conserved Hamiltonian and particle number (L^2 norm squared) for finite-dimensional spectral truncations of the NLS dynamics. The continuum limits of these approximated statistical equilibrium ensembles on finite-dimensional phase spaces are analyzed, holding the energy and particle number at fixed, finite values. The analysis shows that the coherent structure minimizes total energy for a given value of particle number and hence is a solution to the NLS ground state equation, and that the remaining energy resides in Gaussian fluctuations equipartitioned over wavenumbers. Some results of direct numerical integration of the NLS equation are included to validate empirically these properties of the most probable states for the statistical model. Moreover, a theoretical justification of the mean–field approximation is given, in which the approximate ensembles are shown to concentrate on the associated microcanonical ensemble in the continuum limit.

PACS: 05.20.–y; 05.45.–a; 52.35.Mw
Keywords: Nonlinear Schrödinger equation; Coherent structures; Statistical equilibria; Mean–field theory.

Submitted to **Physica D**

1 Introduction

The appearance of macroscopic organized states, or coherent structures, in the midst of turbulent small-scale fluctuations is a common feature of many fluid and plasma systems [1]. Perhaps the most familiar example is the formation of large-scale vortex structures in a turbulent, large Reynolds' number, two-dimensional fluid [2, 3]. A similar phenomenon occurs in slightly dissipative magnetofluids in two and three dimensions, where coherent structures emerge in the form of magnetic islands with flow [4, 5]. In the present work, we shall be concerned with another nonlinear partial differential equation whose solutions in the long-time limit tend to form large-scale coherent structures while simultaneously exhibiting intricate fluctuations on very fine spatial

*Corresponding author: phone: (505) 667-8715; fax: (505) 665-2659; e-mail: rjordan@cnls.lanl.gov

scales. Namely, we shall consider the class of one-dimensional wave systems governed by a nonlinear Schrödinger equation of the form

$$i\psi_t + \psi_{xx} + f(|\psi|^2)\psi = 0\,, \tag{1}$$

for various nonlinearities f chosen so that the dynamics is nonintegrable and free of wave collapse.

Numerical simulations of the NLS equation (1) in a bounded spatial interval with either periodic or Dirichlet boundary conditions demonstrate that, after a sufficiently long time, the field ψ evolves into a state consisting of a time-periodic coherent structure on the large spatial scales coupled with turbulent fluctuations, or radiation, on the small scales [6, 7, 8]. For a focusing nonlinearity f ($f(a) > 0, f'(a) > 0$ for $a > 0$) the coherent structure takes the form of a spatially localized, solitary wave. This organization into a soliton-like state is observed from generic initial conditions for the focusing NLS equation [8]. At intermediate times a typical solution consists of a collection of the soliton-like structures, which as time evolves undergo a succession of collisions, or interactions. In these collisions the larger of the solitons increases in amplitude and smaller waves are radiated. This interaction of the solitons continues until eventually a single soliton of large amplitude survives in a sea of small-scale turbulent waves. (Even though (1) does not possess solitons in the strict sense unless $f(|\psi|^2) = |\psi|^2$, it does have spatially localized solitary wave solutions in the generic focusing case. As a matter of convenience, and in keeping with the terminology in [6, 7], we therefore simply refer to the coherent states that emerge in the simulations in this case as solitons). Figure (1) below illustrates this behavior for the particular saturated focusing nonlinearity $f(|\psi|^2) = |\psi|^2/(1+|\psi|^2)$. For a defocusing f, coherent structures also emerge and persist, but they are less conspicuous since they are not spatially localized. The reader is referred to [8] for a detailed description of numerical investigations of the long-time behavior of the system (1) for various nonlinearities f.

Theoretical arguments and numerical experiments suggest that these coherent structures for the NLS dynamics (1) correspond to ground states of the system. That is, a coherent structure assumes the form $\psi(x,t) = \phi(x)\exp(-i\lambda t)$, where $\phi(x)$ is a solution of the ground state equation

$$\phi_{xx} + f(|\phi|^2)\phi + \lambda\phi = 0\,, \tag{2}$$

In turn, these ground states ϕ are the minimizers of the Hamiltonian

$$H(\psi) = \frac{1}{2}\int_\Omega |\psi_x|^2\,dx - \frac{1}{2}\int_\Omega F(|\psi|^2)\,dx \tag{3}$$

subject to the constraint that the particle number

$$N(\psi) = \frac{1}{2}\int_\Omega |\psi|^2\,dx \tag{4}$$

be equal to given initial value N^0. We shall refer to the first and second terms in (3) as the *kinetic energy* and the *potential energy*, respectively. The potential F is defined by $F(a) = \int_0^a f(a')da'$. The parameter λ arising in (2) is the Lagrange multiplier in the variational principle: minimize $H(\phi)$ subject to $N(\phi) = N^0$. For focusing nonlinearities f, the ground states ϕ are spatially localized. By contrast, for periodic boundary conditions and for defocusing nonlinearities f, the ground states are spatially uniform; this can be proved by a straightforward application of Jenson's inequality when the potential F is strictly concave. Nevertheless, independent of the properties of the ground states, the coherent structures are expected to be constrained minimizers of the Hamiltonian given the particle number, whenever these minimizers exist.

Yankov and collaborators [9, 6, 7] have suggested that the tendency of the solution of the NLS system in the focusing case to approach a coherent soliton state coupled with small-scale radiation can be understood within a statistical mechanics framework. They argue that the solutions to the ground state equation that realize the minimum of the Hamiltonian subject to the constraint on the particle number (2) are "statistical attractors" to which the solutions of the NLS equation (1) tend to relax. In this argument, the process which increases the amplitude of the solitons as the number of solitons decreases is thought to be "thermodynamically advantageous", in the sense that it increases the "entropy" of the system. This entropy seems to be directly related to the amount of kinetic energy contained in the radiation, or the small-scale fluctuations of the field

ψ, so that the transference of kinetic energy to the fluctuations is accompanied by an increase in the entropy. Similar ideas have been expounded by Pomeau [10] who has argued, based on weak turbulence theory, that the defocusing cubic NLS equation ($f(|\psi|^2) = -|\psi|^2$) in a bounded two–dimensional spatial domain should exhibit the tendency to approach a long-time state consisting of solution of the ground state equation (2) plus small-scale radiation.

Our primary purpose in the present paper is to construct a statistical equilibrium model of the coherent structures and the turbulent fluctuations inherent in the long–time behavior of system (1). With this model we seek to translate the various intuitive ideas about the balance between order and disorder in the NLS dynamics into explicit, verifiable calculations. In particular, we provide strong support to the notion set forth in [9, 6, 7, 10] that the ground states, which minimize the energy for a given particle number, are statistically preferred states. In addition, we furnish a definite meaning to the concept that random fluctuations absorb the remainder of the energy.

We derive our equilibrium statistical model from a mean–field approximation of the conserved Hamiltonian H and particle number N for a finite–dimensional truncation of the NLS equation (1). This approximation relies on the fact that the fluctuations of ψ about its mean, the coherent structure, become asymptotically small in the continuum limit, provided that the mean energy and mean particle number remain finite as the number of modes in the spectral truncations goes to infinity. In constructing the mean–field theory, therefore, we choose that probability measure on the finite–dimensional phase space which maximizes entropy subject to approximated constraints on the mean energy and particle number. Specifically, we approximate the means of the particle number and the potential energy term in the Hamiltonian in terms of the mean state $\langle\psi\rangle$.

This approach gives rise to a Gaussian model, which is accordingly easy to analyze. Thus, we find that to any initial value N^0 of the particle number there corresponds a coherent structure, which is precisely the mean state, or most probable state, for the statistical model. Moreover, we show that this mean state is indeed a solution to the ground state equation that minimizes the Hamiltonian over all states with the same particle number N^0. Furthermore, we clarify the role of fine-scale fluctuations by demonstrating that the gradient field ψ_x has finite variance at every point. Thus, we see that the difference between the initial energy H^0 and the energy of the mean state, $H(\langle\psi\rangle)$, resides entirely in the kinetic energy of infinitesimally fine-scale fluctuations. This result, which makes precise the ideas of Yankov *et al.* [6, 7] and Pomeau [10], explains how it is possible for an ergodic solution to the Hamiltonian system (1) to approach a solution to the ground state equation (2) without violating the conservation of energy or particle number.

In the present paper the principal justification for our mean–field approximation is the evidence of extensive numerical simulations, which show that the fluctuations in ψ go to zero over longer and longer time intervals [6, 7, 8]. Nevertheless, we also include an *a posteriori* justification of this approximation by proving that, in the continuum limit, the probability measures defining the mean–field theory with approximately conserved quantities H and N concentrate about the phase space manifold on which $H = H^0$ and $N = N^0$. In other words, we establish a form of the equivalence of ensembles, showing that our mean-field ensemble becomes equivalent in a certain sense to the microcanonical ensemble, which is a measure concentrated on the manifold $H = H^0, N = N^0$. Besides connecting our mean-field approach to the standard principles of equilibrium statistical mechanics [11], this result shows that the mean–field theory is asymptotically exact in the continuum limit. This property of the mean–field ensembles is a common feature of turbulent continuum systems in which the mean values of conserved quantities are fixed while the number of degrees of freedom goes to infinity [12, 13].

These topics are taken up after a quick review of NLS theory and previous work on invariant measures for NLS.

2 The NLS Equation and its Conserved Quantities

The NLS equation (1) describes the slowly-varying envelope of a wave train in a dispersive conservative system and so arises in many branches of physics. It models, for example, gravity waves on deep water [14], Langmuir waves in plasmas [15], pulse propagation along optical fibers [16], and self-induced motion of vortex filaments [17, 18]. In particular, the cubic NLS equation, which corresponds to the nonlinearity $f(|\psi|^2) = \pm|\psi|^2$, has

garnered much attention in the mathematics and physics literature. On the whole real line, or on a periodic interval, the cubic NLS equation is completely integrable via the inverse scattering transform [19, 20]. The NLS equation with any other nonlinearity or boundary conditions, however, is not known to be integrable.

In the present article we are interested in the NLS equation (1) as a generic model of nonlinear wave turbulence. For this reason, we wish to consider nonlinearities f and boundary conditions such that the dynamics is well-posed for all time and yet is nonintegrable. For instance, a natural class of prototypes of this kind consists of those NLS equations with bounded nonlinearities f on a bounded spatial interval with Dirichlet boundary conditions. Bounded nonlinearities arising in physical applications are often called saturated nonlinearities, of which $f(|\psi|^2) = |\psi|^2/(1+|\psi|^2)$ and $f(|\psi|^2) = 1 - \exp(-|\psi|^2)$ are examples. These amount to corrections to the focusing cubic nonlinearity for large wave amplitudes, and they have been proposed as a models of nonlinear self-focusing of laser beams [21], propagation of nonlinear optical waves through dielectric and metallic layered structures [22], and self-focusing of cylindrical light beams in plasmas due to the ponderomotive force [23]. Such nonlinearities certainly satisfy the requirements set forth by Zhidkov to guarantee well–posedness of the NLS equation (1) in the space L^2 for Dirichlet boundary conditions [24].

In light of these properties, we shall adopt the NLS equation with homogeneous Dirichlet boundary conditions on a bounded interval $\Omega = [0, L]$ and with a (focusing or defocusing) bounded nonlinearity f as the generic system throughout our main development. Precisely, we impose on f is the following *boundedness condition*

$$\sup_{a \in R} \left(|f(a)| + |(1+a)f'(a)| \right) \, < \, \infty \, . \tag{5}$$

In Section 9 we indicate how our main results can be extended to a natural class of unbounded nonlinearities. Also, when it is convenient for comparison with some numerical simulations, we can instead use periodic boundary conditions. The necessary modifications, which are straightforward, are left to the reader.

The NLS equation (1) may be cast in the Hamiltonian form

$$\frac{i}{2}\psi_t = \frac{\delta H}{\delta \psi^*} \, ,$$

where ψ^* is the complex conjugate of the field ψ and H is the Hamiltonian defined by (3). For our class of generic NLS systems, the total energy H and the particle number N, or L^2 norm squared, are the only known dynamical invariants for (1). These invariants play the leading role in the statistical equilibrium theory and the values H^0 and N^0, derived from a given initial state, say, alone determine the statistical ensemble.

For later use, we write the NLS equation (1) as the following coupled system of partial differential equations for the real and imaginary components $u(x,t)$ and $v(x,t)$ of the complex field $\psi(x,t)$:

$$u_t + v_{xx} + f(u^2+v^2)v = 0 \, , \quad v_t - u_{xx} - f(u^2+v^2)u = 0 \, . \tag{6}$$

Similarly, the Hamiltonian can be written in terms of the fields u and v as

$$H(u,v) = K(u,v) \, + \, \Theta(u,v) \tag{7}$$

with kinetic and potential energy terms

$$K(u,v) = \frac{1}{2}\int_\Omega (u_x^2 + v_x^2)\, dx \, , \qquad \Theta(u,v) = -\frac{1}{2}\int_\Omega F(u^2+v^2)\, dx \, ,$$

and the particle number is given by

$$N(u,v) = \frac{1}{2}\int_\Omega (u^2+v^2)\, dx \, . \tag{8}$$

In the same way, the ground state equation (2) for $\phi = u + iv$ can be expressed as the following coupled system of differential equations:

$$u_{xx} + f(u^2+v^2)u + \lambda u = 0 \, , \quad v_{xx} + f(u^2+v^2)v + \lambda v = 0 \, . \tag{9}$$

3 Invariant Measures for NLS

Several authors have constructed invariant measures for the NLS equation (1). Zhidkov [24] considered nonlinearities f satisfying the boundedness condition (5) and showed that the Gibbs measure

$$P_\beta(dudv) = Z^{-1} \exp(-\beta H(u,v)) \prod_{x\in\Omega} du(x)dv(x)\,, \tag{10}$$

for fixed $\beta > 0$ is normalizable and has a rigorous interpretation as an invariant measure for the NLS system on the phase space L^2. Similar results have been obtained by Bidegary [25]. These results hold for both periodic and Dirichlet boundary conditions. However, P_β as constructed by Zhidkov fails to account for the invariance of the particle number (4) under the NLS dynamics.

Earlier, Lebowitz *et al.* [26] considered the problem of constructing invariant measures for NLS on the periodic interval with the focusing power law nonlinearities $f(|\psi|^2) = |\psi|^s$. For such nonlinearities the Hamiltonian H is not bounded below, and therefore the measure P_β defined by (10) is not normalizable. As an alternative, Lebowitz *et al.* conditioned the Gibbs measure on the particle number and studied the measures

$$P_{\beta,N^0}(dudv) = Z^{-1} \exp(-\beta H(u,v)) I(N(u,v) \le N^0) \prod_{x\in\Omega} du(x)dv(x)\,, \tag{11}$$

where $I(A)$ is the characteristic function of the set A. They showed that P_{β,N^0} is normalizable for $1 \le s < 4$ for any N^0, and for $s = 4$ if N^0 is sufficiently small. For $s > 4$, though, P_{β,N^0} fails to be normalizable.

The normalizability of the conditioned Gibbs ensemble (11) is closely related to the well-posedness of the NLS equation as an initial value problem [26]. Indeed, for the focusing power law nonlinearities it can be shown that the NLS equation (1) with periodic or Dirichlet boundary conditions is well-posed, or free of wave collapse, in the Sobolev space $H^1(\Omega)$ as long as $s < 4$, while blow-up can occur from smooth initial conditions when $s \ge 4$ [26, 27, 28]. Bourgain [27] proved that under the conditions on s and N^0 considered by Lebowitz *et al.* the measures P_{β,N^0} are rigorously invariant under the NLS dynamics and he thereby obtained a new global existence theorem for the initial value problem with these nonlinearities. Analogous results have also been established by McKean [29].

A common feature of all these invariant measures for the NLS equation is that for any fixed $\beta > 0$ the kinetic energy $(1/2)\int(u_x^2 + v_x^2)\,dx$ is infinite with probability 1, and hence the Hamiltonian is also infinite with probability 1 [26]. This difficulty occurs whenever classical Gibbs statistical mechanics is applied to a system with infinitely many degrees of freedom and is related to the well-known Jeans ultraviolet catastrophe [11]. Consequently, these ensembles are appropriate when the typical state of the continuum system is such that each of its modes contains a finite energy. They are not appropriate, on the other hand, to situations in which a typical state of the system is realized by an ergodic evolution from an ensemble of initial conditions having finite energy H^0 and particle number N^0, both of which partition amongst the many modes of the system.

Since the purpose of the current paper is to construct ensembles which provide meaningful predictions about the long–time behavior of solutions of the NLS equation (1), we are interested in statistical equilibrium ensembles which, in the continuum limit, have finite mean energy and mean particle number. In view of this goal, we scale the inverse temperature parameter β in the Gibbs measure with the number of degrees of freedom n and thereby maintain the mean energy at a finite value as $n \to \infty$. The continuum limit we obtain under this scaling is degenerate in certain ways, but this is hardly surprising given that the known invariant measures for NLS are supported on fields with infinite kinetic energy. On the other hand, the analysis of the scaled continuum limit is greatly simplified by the fact that the random fluctuations of (u, v) at each point tend to zero. This fortuitous property implies that a natural mean–field approximation is asymptotically exact. Moreover, the mean-field ensemble concentrates on the microcanonical ensemble for the invariants H and N in the scaled continuum limit. For this reason, we can avoid the canonical ensemble constructed from H and N, which fails to be normalizable for focusing nonlinearities because the linear combination $H - \lambda N$ always has a direction in which it goes to $-\infty$, as we shall demonstrate in Section 5.

4 Discretization of NLS

We now introduce a finite–dimensional approximation of the NLS equation (1) with homogeneous Dirichlet boundary conditions on the interval $\Omega = [0, L]$. For this purpose we shall use a standard spectral truncation, even though many other approximation schemes could be invoked. Let $e_k(x) = \sqrt{2/L}\sin(\sqrt{\lambda_k}x)$ and $\lambda_k = (k\pi/L)^2, k = 1, 2, \ldots$, be the eigenfunctions and eigenvalues of the operator $-\frac{d^2}{dx^2}$ on Ω with homogeneous Dirichlet boundary conditions. The eigenfunctions e_k form an orthonormal basis for the real space $L^2(\Omega)$. For given n, let $X_n = \text{span}\,\{e_1, \ldots, e_n\} \subset L^2$, define $u^{(n)}$ and $v^{(n)}$ in X_n by

$$u^{(n)}(x,t) = \sum_{k=1}^{n} u_k(t)e_k(x)\,, \quad v^{(n)}(x,t) = \sum_{k=1}^{n} v_k(t)e_k(x)\,, \tag{12}$$

and set $\psi^{(n)}(x,t) = u^{(n)}(x,t) + iv^{(n)}(x,t)$. We consider the following evolution equation for $\psi^{(n)}$:

$$i\psi_t^{(n)} + \psi_{xx}^{(n)} + P^{(n)}\left(f(|\psi^{(n)}|^2)\psi^{(n)}\right) = 0\,, \tag{13}$$

where $P^{(n)}$ denotes the orthogonal projection from $L^2(\Omega)$ onto X_n. Equation (13) is clearly a spectral truncation of the NLS equation (1). It is equivalent to the following coupled system of ordinary differential equations for the Fourier coefficients u_k and $v_k, k = 1, \ldots, n$:

$$\frac{du_k}{dt} - \lambda_k v_k + \int_\Omega f((u^{(n)})^2 + (v^{(n)})^2)v^{(n)}e_k\,dx = 0\,, \tag{14}$$

$$\frac{dv_k}{dt} + \lambda_k u_k - \int_\Omega f((u^{(n)})^2 + (v^{(n)})^2)u^{(n)}e_k\,dx = 0, \tag{15}$$

where we have used the fact that $(e_k)_{xx} = -\lambda_k e_k$.

Whether viewed as the evolution of $(u^{(n)}, v^{(n)})$ on $X_n \times X_n$ or in terms of the Fourier coefficients (u_k, v_k) on R^{2n}, it can be shown [24, 25] that this finite–dimensional dynamical system has Hamiltonian structure, with Hamiltonian $H_n = K_n + \Theta_n$, where

$$K_n(u^{(n)}, v^{(n)}) = \frac{1}{2}\int_\Omega [(u_x^{(n)})^2 + (v_x^{(n)})^2]\,dx = \frac{1}{2}\sum_{k=1}^{n}\lambda_k(u_k^2 + v_k^2)\,, \tag{16}$$

is the kinetic energy, and

$$\Theta_n(u^{(n)}, v^{(n)}) = -\frac{1}{2}\int_\Omega F((u^{(n)})^2 + (v^{(n)})^2)\,dx\,, \tag{17}$$

is the potential energy. The functionals H_n, K_n and Θ_n are just the restrictions to $X_n \times X_n$ of the functionals H, K and Θ, which are defined by (7). The Hamiltonian H_n is, of course, an invariant of the dynamics. The particle number

$$N_n(u^{(n)}, v^{(n)}) = \frac{1}{2}\int_\Omega [(u^{(n)})^2 + (v^{(n)})^2]\,dx = \frac{1}{2}\sum_{k=1}^{n}(u_k^2 + v_k^2) \tag{18}$$

is also conserved by this dynamics, as may be verified by direct calculation. Also, N_n is the restriction of the functional N, which is defined by (8), to the space $X_n \times X_n$.

The Hamiltonian system (14)-(15) as a dynamical system on R^{2n} satisfies the Liouville property [25]. In other words, the Lebesgue measure $\prod_{k=1}^{n} du_k dv_k$ on R^{2n} is invariant under the phase flow for this dynamics.

5 Mean-Field Ensembles for NLS

We now proceed to construct a statistical model that describes the long–time behavior of the truncated NLS system (14)-(15). According to the fundamental principles of equilibrium statistical mechanics, the Liouville

property and the ergodicity of the dynamics provide the usual starting point for such a description, and the microcanonical ensemble is the appropriate statistical distribution (that is, probability measure on phase space) with which to calculate averages for an isolated Hamiltonian system [11, 30]. In the specific system under consideration, the natural canonical random variables are u_k and v_k, which comprise the phase space R^{2n}. The random fields $u^{(n)}(x) = \sum u_k e_k(x)$, $v^{(n)}(x) = \sum v_k e_k(x)$, and $\psi^{(n)}(x) = u^{(n)}(x) + iv^{(n)}(x)$, as well as the functions $H_n = H(u^{(n)}, v^{(n)})$ and $N_n = N(u^{(n)}, v^{(n)})$, are determined by these canonical variables. The microcanonical ensemble for the truncated NLS system therefore takes the form

$$P_{H^0,N^0}(dudv) \;=\; W^{-1}\delta(H_n - H^0)\,\delta(N_n - N^0)\prod_{k=1}^{n} du_k dv_k\,, \tag{19}$$

where $W = W(H^0, N^0)$ is the structure function or normalizing factor. Under the ergodic hypothesis, expectations with respect to this ensemble equal long-time averages over the spectrally-truncated dynamics from initial conditions with prescribed, finite values H^0 and N^0 of the invariants H_n and N_n.

While the microcanonical ensemble is a well-defined invariant measure for the spectrally-truncated dynamics, it is a cumbersome to analyze, and the calculation of ensemble averages for finite n is not feasible. In statistical mechanics the usual procedure to overcome this difficulty is to introduce another invariant measure, the canonical ensemble, and prove that in the limit as $n \to \infty$, it becomes equivalent to the microcanonical ensemble [11, 30]. For the NLS system the canonical Gibbs ensemble is

$$P_{\beta,\lambda}(dudv) \;=\; Z^{-1}\exp(-\beta[H_n - \lambda N_n])\prod_{k=1}^{n} du_k dv_k, \tag{20}$$

where β and λ are chosen such that the averages of H_n and N_n with respect to $P_{\beta,\lambda}$ equal H^0 and N^0, respectively, and $Z = Z(\beta, \lambda)$ is the partition function or normalizing factor. Unfortunately, for general focusing nonlinearities f, even those satisfying the boundedness condition (5), this canonical measure is not normalizable. To see this degeneracy, let us consider a nonlinearity f that is both positive and strictly increasing and an associated ground state ϕ that is a nontrivial solution to (2). Then, for any positive scale factor σ, a straightforward calculation yields the identity

$$\begin{aligned}(H - \lambda N)(\sigma\phi) \;&=\; \frac{\sigma^2}{2}\int_\Omega [|\phi_x|^2 \,-\, \lambda|\phi|^2]\,dx \,-\, \frac{1}{2}\int_\Omega F(\sigma^2|\phi|^2)\,dx \\ &=\; -\frac{1}{2}\int_\Omega dx\int_0^{\sigma^2|\phi|^2}[f(a) - f(|\phi|^2)]\,da\,.\end{aligned}$$

It follows from this expression that, as σ increases to infinity, the value of $H - \lambda N$ at $\sigma\phi$ tends to $-\infty$. The same reasoning applies to the discretized system. Consequently, the function $H_n - \lambda N_n$ in the Gibbs weight has a direction in which is goes to $-\infty$, resulting in the divergence of the partition function Z in (20). In essence, this divergence reflects the fact that such a ground state ϕ is a critical point, but not a minimizer, of $H - \lambda N$. Locally, the first variation $\delta H - \lambda\delta N$ vanishes at ϕ, but the second variation $\delta^2 H - \lambda\delta^2 N$ is positive only on those variations $\delta\phi$ for which $\delta N = 0$. When $\delta\phi = (1 - \sigma)\phi$ for σ near 1, and thus $\delta N \neq 0$, the corresponding second variation of $H - \lambda N$ is negative. The above calculation shows that this local behavior extends to a global degeneracy for a wide class of nonlinearities. As Lebowitz *et al.* discuss in [26], this defect of the canonical ensemble is even more severe when the nonlinearity f is unbounded.

These considerations compel us to find another way to approximate the microcanonical ensemble and thus to derive a tractable statistical equilibrium model. First, we note that since the microcanonical measure (19) depends on $u^{(n)}$ and $v^{(n)}$ through H_n and N_n only, it is invariant with respect to multiplicative constants of the form $e^{i\theta}$. In view of this phase invariance, we can factor the random field $\psi^{(n)}$ in the form $e^{i\theta}\phi^{(n)}$, where θ is uniformly distributed on $[0, 2\pi]$ and is independent of $\phi^{(n)}$. In what follows, we shall write $\phi^{(n)} = u^{(n)} + iv^{(n)}$ with the phase normalization

$$\arg\int_\Omega \phi^{(n)}(x)\,dx \;=\; 0\,,$$

and consider the microcanonical distribution conditioned by this constraint. It is this conditional distribution that we seek to approximate with a mean-field approximation. In constructing the approximation, we can therefore describe distributions for $\phi^{(n)}$ that need not possess invariance under a change of phase, with the understanding that we can recover the random field $\psi^{(n)}$ by setting $\psi^{(n)} = e^{i\theta}\phi^{(n)}$.

The key to the mean-field construction is supplied by direct numerical simulations of the evolving microstates ψ governed by (1) [7, 8]. Since the particle number and the Hamiltonian are well conserved in these simulations, we may think of these numerical experiments as realizations of the microcanonical ensemble. The simulations clearly demonstrate that the amplitude of fluctuations of the random field ϕ (the phase normalized field associated with ψ) at each point $x \in \Omega$ become small in the long-time limit, and that they contribute little to the conserved L^2 norm of ϕ [7, 8] (also, see Figure 1 below). Moreover, this effect becomes more apparent as the spatial resolution of the numerical simulations is improved [8]. In particular, these fluctuations become negligible compared to the magnitude of the coherent structure that emerges in the mean field ϕ, which in the focusing cases takes the form of a single, localized soliton. On the basis of these properties of the simulated microstates, we shall adopt the hypothesis that, in the limit of infinite resolution, the fluctuations of ϕ are infinitesimal. Precisely, we shall make the following *vanishing of fluctuations hypothesis*:

With respect to the phase-invariance conditioned microcanonical ensemble on $2n$-dimensional phase-space, there holds

$$\lim_{n\to\infty} \int_\Omega [\mathrm{Var}(u^{(n)}(x)) + \mathrm{Var}(v^{(n)}(x))]\, dx \;=\; \lim_{n\to\infty} \sum_{k=1}^{n} [\mathrm{Var}(u_k) + \mathrm{Var}(v_k)] \;=\; 0\,, \tag{21}$$

where Var denotes the variance of the indicated random variable with respect to that ensemble. The first equality follows from the identity $\int_\Omega [\mathrm{Var}(u^{(n)}(x)) + \mathrm{Var}(v^{(n)}(x))]\, dx = \sum_{k=1}^{n} [\mathrm{Var}(u_k) + \mathrm{Var}(v_k)]$.

Our strategy is to use the vanishing of fluctuations hypothesis to build a mean-field ensemble that is equivalent in an appropriate sense to the microcanonical ensemble (19) in the continuum limit as $n \to \infty$, but is easy to analyze for finite n. Rather than attempt to verify this hypothesis *a priori*, we shall show *a posteriori* that the resulting mean-field ensemble concentrates on the microcanonical manifold at fixed H^0 and N^0. The analysis of the fundamental microcanonical ensemble itself will be taken up in a subsequent publication.

We note however that the numerical simulations of the NLS system which support the hypothesis (21) also strongly indicate that the fluctuations of the gradients $u_x^{(n)}, v_x^{(n)}$ are not negligible in the limit of infinite resolution. Figure 2 illustrates the typical long-time behavior of $|\psi_x|^2$ for the particular nonlinearity $f(|\psi|^2) = |\psi|^2/(1+|\psi|^2)$, in which finite fluctuations in the gradient ψ_x persist. This property of the microcanonical ensemble is compatible with the vanishing of fluctuations hypothesis, as can be seen from the straightforward identity

$$\int_\Omega [\mathrm{Var}(u_x^{(n)}(x)) + \mathrm{Var}(v_x^{(n)}(x))] dx = \sum_{k=1}^{n} \lambda_k [\mathrm{Var}(u_k) + \mathrm{Var}(v_k)]\,. \tag{22}$$

In fact, as we shall see, this quantity remains finite in the continuum limit according to our statistical theory, and it represents the portion of kinetic energy absorbed by infinitesimally fine-scale fluctuations.

We now proceed to construct a mean-field ensemble for each n based on approximations derived from the hypothesis (21). This ensemble is determined by a probability density $\rho^{(n)}$ on R^{2n} with respect to Lebesgue measure $\prod_{k=1}^{n} du_k dv_k$. (Unlike the microcanonical ensembles, the mean-field ensembles are taken to be absolutely continuous with respect to Lebesgue measure for each n.) In order to define that $\rho^{(n)}$ which governs the mean-field theory, we appeal to standard statistical-mechanical and information-theoretic principles [31, 11, 30] and choose $\rho^{(n)}$ so that it maximizes the Gibbs–Boltzmann entropy functional

$$S(\rho) = -\int_{R^{2n}} \rho(u_1,\ldots,u_n,v_1,\ldots,v_n) \log \rho(u_1,\ldots,u_n,v_1,\ldots,v_n) \prod_{k=1}^{n} du_k dv_k\,, \tag{23}$$

subject to some appropriate constraints. The entropy S has the well-known interpretation as a measure of the number of *microstates* (u, v) corresponding to the *macrostate* ρ [11]. The form of the entropy as a functional of

ρ is dictated by the Liouville property of the dynamics (14)–(15), which requires that the entropy be relative to the invariant measure $\prod_{k=1}^{n} du_k dv_k$. Alternatively, from the information theoretic point of view, maximizing the entropy S amounts to finding the least biased distribution on $2n$-dimensional phase space compatible with the constraints and the uniform prior distribution $\prod_{k=1}^{n} du_k dv_k$ [31].

The constraints imposed in the Maximum Entropy Principle (MEP) are the crucial ingredient in the determination of the ensemble $\rho^{(n)}$. As is well-known, the microcanonical ensemble is produced when they are taken to be the exact constraints $H_n = H^0$ and $N_n = N^0$; the canonical ensemble, which however is ill-defined in the focusing case, corresponds to average constraints $\langle H_n \rangle = H^0$ and $\langle N_n \rangle = N^0$. Throughout our discussion of (MEP) the angle brackets $\langle \cdot \rangle$ denote expectation with respect to the admissible density ρ in (MEP). For our mean-field theory we seek constraints that are intermediate between those giving the microcanonical and the canonical ensembles. To this end we invoke the vanishing of fluctuations hypothesis (21) and impose the following constraints:

$$\frac{1}{2}\sum_{k=1}^{n}(\langle u_k \rangle^2 + \langle v_k \rangle^2) = N^0 \,. \tag{24}$$

$$\frac{1}{2}\sum_{k=1}^{n}\lambda_k(\langle u_k^2 \rangle + \langle v_k^2 \rangle) - \frac{1}{2}\int_\Omega F(\langle u^{(n)} \rangle^2 + \langle v^{(n)} \rangle^2)\, dx = H^0, \tag{25}$$

which we shall refer to as the *mean–field constraints* for obvious reasons. We note that these constraints involve only the first and second moments of the fields $u^{(n)}$ and $v^{(n)}$ with respect to ρ.

We can motivate this choice of constraints in (MEP) by approximating the values of the conserved quantities H_n and N_n in terms of means with respect to the microcanonical ensemble. An immediate implication of (21) applied to the definition of N_n given in (18) is that

$$\begin{aligned} N^0 &= \frac{1}{2}\int_\Omega (\langle u^{(n)}(x) \rangle^2 + \langle v^{(n)}(x) \rangle^2)\, dx \;+\; \frac{1}{2}\int_\Omega [\mathrm{Var}(u^{(n)}(x)) + \mathrm{Var}(v^{(n)}(x))]\, dx \\ &= N_n(\langle u^{(n)} \rangle, \langle v^{(n)} \rangle) \;+\; o(1) \end{aligned} \tag{26}$$

as $n \to \infty$. When we drop the error term in this approximation at finite n and we replace expectation with respect to the microcanonical distribution by expectation with respect to an admissible density ρ, we obtain the mean-field constraint (24).

Similarly, we can obtain the mean-field constraint (25) from the vanishing of fluctuations hypothesis (21) by analyzing H_n with respect to the microcanonical distribution for large n. The kinetic energy (16) in $H^0 = \langle H_n \rangle$ is retained exactly as $\langle K_n \rangle$. On the other hand, the potential energy (17) is approximated by expanding F about the mean $(\langle u^{(n)} \rangle, \langle v^{(n)} \rangle)$, which yields

$$\begin{aligned} \Theta_n(u^{(n)}, v^{(n)}) &= -\frac{1}{2}\int_\Omega F(\langle u^{(n)} \rangle^2 + \langle v^{(n)} \rangle^2)\, dx \\ &- \int_\Omega f(\langle u^{(n)} \rangle^2 + \langle v^{(n)} \rangle^2)\left(\langle u^{(n)} \rangle (u^{(n)} - \langle u^{(n)} \rangle) + \langle v^{(n)} \rangle (v^{(n)} - \langle v^{(n)} \rangle)\right) dx \\ &- \frac{1}{4}\int_\Omega \begin{pmatrix} (u^{(n)} - \langle u^{(n)} \rangle) \\ (v^{(n)} - \langle v^{(n)} \rangle) \end{pmatrix}^T J(\tilde{u}^{(n)}, \tilde{v}^{(n)}) \begin{pmatrix} (u^{(n)} - \langle u^{(n)} \rangle) \\ (v^{(n)} - \langle v^{(n)} \rangle) \end{pmatrix} dx \,, \end{aligned} \tag{27}$$

where

$$J(u,v) = \begin{pmatrix} 2f(u^2+v^2) + 4u^2 f'(u^2+v^2) & 4uvf'(u^2+v^2) \\ 4uvf'(u^2+v^2) & 2f(u^2+v^2) + 4v^2 f'(u^2+v^2) \end{pmatrix} \tag{28}$$

is the matrix of second partial derivatives of $F(u^2+v^2)$ with respect to u and v and $(\tilde{u}^{(n)}, \tilde{v}^{(n)})$ lies between $(u^{(n)}, v^{(n)})$ and $(\langle u^{(n)} \rangle, \langle v^{(n)} \rangle)$. Because we are considering nonlinearities f satisfying (5) each element of the matrix $J(\tilde{u}^{(n)}, \tilde{v}^{(n)})$ is bounded over the domain Ω independently of n. Therefore, taking the expectation on

both sides of equation (27) and noting that the second term on the right hand side of this equation has zero mean, we obtain

$$\langle \Theta_n \rangle = -\frac{1}{2} \int_\Omega F(\langle u^{(n)} \rangle^2 + \langle v^{(n)} \rangle^2) \, dx + R_n \, ,$$

where the remainder R_n satisfies

$$\begin{aligned} |R_n| &\leq C \left\langle \int_\Omega [(u^{(n)} - \langle u^{(n)} \rangle)^2 + (v^{(n)} - \langle v^{(n)} \rangle)^2] \, dx \right\rangle \\ &= C \int_\Omega [\mathrm{Var}(u^{(n)}(x)) + \mathrm{Var}(v^{(n)}(x))] \, dx \end{aligned}$$

for some constant C independent of n. The vanishing of fluctuations hypothesis (21) thus implies that $R_n \to 0$ as $n \to \infty$, and hence that

$$H^0 \; = \; \frac{1}{2} \sum_{k=1}^{n} \lambda_k (\langle u_k^2 \rangle + \langle v_k^2 \rangle) - \frac{1}{2} \int_\Omega F(\langle u^{(n)} \rangle^2 + \langle v^{(n)} \rangle^2) \, dx \; + \; o(1). \tag{29}$$

The mean-field constraint (25) results from neglecting the error term in this expression for finite n and replacing microcanonical expectations by expectations with respect to ρ.

In summary, the mean-field theory is defined by a probability density $\rho^{(n)}$ on R^{2n} that maximizes the entropy S given in (23) subject to the constraints (24) and (25). These mean-field ensembles, which solve the governing (MEP), are analyzed in detail in the subsequent sections. Here we merely note that they enjoy some properties not shared by either the microcanonical or the canonical ensembles. First, unlike the microcanonical ensemble, the mean-field ensembles are analytically tractable because the densities $\rho^{(n)}$ are Gaussian. This desirable property is a simple consequence of the fact that the constraints (24) and (25) involve only the first and second moments of $\rho^{(n)}$. Second, in contrast to the canonical ensemble, the mean-field ensemble exists in both the focusing and defocusing cases. This crucial property depends on the fact that fluctuations in N_n are suppressed in the mean-field ensemble.

However, because of the approximations made in deriving the constraints (24) and (25), the mean-field ensemble $\rho^{(n)}$ is not an invariant measure for the truncated NLS dynamics (14)-(15) at finite n, even with the random phase $e^{i\theta}$ included. Nevertheless, it becomes consistent with the fundamental microcanonical ensemble in the limit as $n \to \infty$, in the sense that $\langle N_n \rangle \to N^0$, $\langle H_n \rangle \to H^0$, and the variances of N_n and H_n converge to 0 . These properties, which are proved in Section 8, imply that the ensembles $\rho^{(n)}$ concentrate about the phase space manifold $H_n = H^0, N_n = N^0$, and thereby justify the approximations made in deriving the mean-field theory.

6 Equilibrium States for the Mean-Field Theory

We now proceed to calculate and analyze the solutions $\rho^{(n)}$ to the maximum entropy principle (MEP). First, we calculate the density $\rho^{(n)}$ which maximizes entropy subject to the mean-field constraints (24)-(25) by the method of Lagrange multipliers. If we denote the left-hand sides of (24) and (25) by $\tilde{N}_n$ and $\tilde{H}_n$, then we have

$$\delta S = \mu^{(n)} \delta \tilde{N}_n + \beta^{(n)} \delta \tilde{H}_n \, , \tag{30}$$

where δ denotes variation with respect to the density variable ρ and $\mu^{(n)}$ and $\beta^{(n)}$ are the Lagrange multipliers to enforce the constraints (24) and (25) on $\tilde{N}_n$ and $\tilde{H}_n$. The variation of S is

$$\delta S = - \int_{R^{2n}} \log(\rho) \delta\rho \prod_{k=1}^{n} du_k dv_k \, , \tag{31}$$

for variations $\delta\rho$ satisfying $\int_{R^{2n}} \delta\rho \prod_{k=1}^{n} du_k dv_k = 0$. Similarly, the variations of $\tilde{N}_n$ and $\tilde{H}_n$ are

$$\delta\tilde{N}_n = \sum_{k=1}^{n} \int_{R^{2n}} (\langle u_k\rangle u_k + \langle v_k\rangle v_k)\, \delta\rho \prod_{k=1}^{n} du_k dv_k\,, \tag{32}$$

and

$$\begin{aligned} \delta\tilde{H}_n = & \sum_{k=1}^{n} \frac{1}{2} \int_{R^{2n}} \lambda_k (u_k^2 + v_k^2)\, \delta\rho \prod_{k=1}^{n} du_k dv_k \\ - & \int_{R^{2n}} \left(\int_\Omega f(\langle u^{(n)}\rangle^2 + \langle v^{(n)}\rangle^2) \left[\langle u^{(n)}\rangle u^{(n)} + \langle v^{(n)}\rangle v^{(n)} \right] dx \right) \delta\rho \prod_{k=1}^{n} du_k dv_k \end{aligned} \tag{33}$$

using the relation $F' = f$. Equating the coefficients of $\delta\rho$ in (30) and recalling that $u^{(n)} = \sum_{k=1}^{n} u_k e_k$ and $v^{(n)} = \sum_{k=1}^{n} v_k e_k$, we discover after some algebraic manipulations that the entropy–maximizing density $\rho^{(n)}$ has the form

$$\rho^{(n)}(u_1, \ldots, u_n, v_1, \ldots, v_n) = \prod_{k=1}^{n} \rho_k(u_k, v_k)\,, \tag{34}$$

where, for $k = 1, \ldots, n$,

$$\rho_k(u_k, v_k) \;=\; \frac{\beta^{(n)}\lambda_k}{2\pi} \exp\left\{ -\frac{\beta^{(n)}\lambda_k}{2} \left[(u_k - \langle u_k\rangle)^2 + (v_k - \langle v_k\rangle)^2 \right] \right\}\,, \tag{35}$$

with

$$\langle u_k\rangle = \frac{1}{\lambda_k} \int_\Omega f(\langle u^{(n)}(x)\rangle^2 + \langle v^{(n)}(x)\rangle^2) \langle u^{(n)}(x)\rangle e_k(x)\, dx - \frac{\mu^{(n)}}{\beta^{(n)}\lambda_k} \langle u_k\rangle\,, \tag{36}$$

$$\langle v_k\rangle = \frac{1}{\lambda_k} \int_\Omega f(\langle u^{(n)}(x)\rangle^2 + \langle v^{(n)}(x)\rangle^2) \langle v^{(n)}(x)\rangle e_k(x)\, dx - \frac{\mu^{(n)}}{\beta^{(n)}\lambda_k} \langle v_k\rangle\,. \tag{37}$$

We see immediately that $u_1, \ldots, u_n, v_1, \ldots, v_n$ are mutually independent Gaussian random variables with means satisfying (36)–(37), and variances

$$\mathrm{Var}(u_k) = \mathrm{Var}(v_k) = \frac{1}{\beta^{(n)}\lambda_k}\,. \tag{38}$$

The multiplier $\beta^{(n)}$ is therefore necessarily positive. The equations (36)–(37) for the means $\langle u_k\rangle$ and $\langle v_k\rangle$ can be written as an equivalent complex equation for the mean field $\langle\phi^{(n)}\rangle = \langle u^{(n)}\rangle + i\langle v^{(n)}\rangle$:

$$\langle\phi^{(n)}\rangle_{xx} \;+\; P^{(n)}\left(f(|\langle\phi^{(n)}\rangle|^2)\langle\phi^{(n)}\rangle \right) \;+\; \lambda^{(n)}\langle\phi^{(n)}\rangle \;=\; 0\,, \tag{39}$$

where we introduce the real parameter $\lambda^{(n)} = -\mu^{(n)}/\beta^{(n)}$. Recalling that $P^{(n)}$ is the projection onto the span X_n of the first n eigenfunctions $e_1, \ldots, e_n$, we recognize this equation as the spectral truncation of the ground state equation (2) for the continuous NLS system (1). In other words, the mean field $\langle\phi^{(n)}\rangle$ at each finite n is a critical point of the functional $H_n - \lambda^{(n)} N_n = 0$, by virtue of (39). Thus, we draw the important conclusion that the mean field predicted by the statistical equilibrium theory is a solution of the ground state equation for the NLS system.

Since the maximum entropy distribution $\rho^{(n)}$ satisfies the mean–field Hamiltonian constraint (25), a direct calculation reveals that

$$\begin{aligned} H^0 &= \frac{1}{2}\sum_{k=1}^{n} \lambda_k (\mathrm{Var}\ u_k + \mathrm{Var}\ v_k) + \frac{1}{2}\sum_{k=1}^{n} \lambda_k (\langle u_k\rangle^2 + \langle v_k\rangle^2) - \frac{1}{2}\int_\Omega F(\langle u^{(n)}\rangle^2 + \langle v^{(n)}\rangle^2)\, dx \\ &= \frac{n}{\beta^{(n)}} + H_n(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)\,, \end{aligned} \tag{40}$$

where (38) is invoked to obtain the second equality. In this expression the term $n/\beta^{(n)}$ in represents the contribution of fluctuations to the energy, while $H_n(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ is the energy of the mean state, or coherent structure. We can rearrange (40) to get the following expression for $\beta^{(n)}$ in terms of the number of modes n and the energy of the mean state:

$$\beta^{(n)} = \frac{n}{H^0 - H_n(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)} \,. \tag{41}$$

Now we come to the following central result, which enables us to characterize completely the mean–field statistical ensembles $\rho^{(n)}$.

Theorem 1. The mean field pair $(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ corresponding to any solution $\rho^{(n)}$ of the constrained variational principle (MEP) is an absolute minimizer of the Hamiltonian H_n over all $(u^{(n)}, v^{(n)}) \in X_n \times X_n$ which satisfy the particle number constraint $N_n(u^{(n)}, v^{(n)}) = N^0$.

Proof: To prove this assertion, we calculate the entropy of a solution $\rho^{(n)}$ of (MEP). Referring to equations (34)–(38), which define any solution of (MEP) we immediately obtain

$$\begin{aligned} S(\rho^{(n)}) &= -\int_{R^{2n}} \rho^{(n)} \log \rho^{(n)} \prod_{k=1}^{n} du_k dv_k \\ &= -\sum_{k=1}^{n} \int_{R^2} \rho_k(u_k, v_k) \log \rho_k(u_k, v_k)\, du_k dv_k \\ &= -\sum_{k=1}^{n} \left(\log \frac{\beta^{(n)}\lambda_k}{2\pi} - 1 \right) \\ &= C(n) + n \log \left(\frac{L^2}{\beta^{(n)}} \right) , \end{aligned}$$

where $C(n) = n - \sum_{k=1}^{n} \log(k^2\pi/2)$ depends only on the number of Fourier modes n. Here, we have used the identity $\lambda_k = (k\pi/L)^2$. From this calculation and equation (41), it follows that

$$S(\rho^{(n)}) = C(n) + n \log \left(\frac{L^2 \left[H^0 - H_n(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)\right]}{n} \right) . \tag{42}$$

Evidently, the entropy $S(\rho^{(n)})$ is maximized if and only if the mean field pair $(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ corresponding to $\rho^{(n)}$ minimizes the Hamiltonian H_n over all fields $(u^{(n)}, v^{(n)}) \in X_n \times X_n$ that satisfy the constraint $N_n(u^{(n)}, v^{(n)}) = N^0$. That such minimizers exist is proved in the Appendix.

The equation (42) has the interesting interpretation that, up to additive and multiplicative constants, the entropy of the mean-field equilibrium $\rho^{(n)}$ is the logarithm of the kinetic energy contained in the fluctuations about the mean state, which is the coherent structure. Because the potential energy is determined entirely by the mean, the difference between the prescribed total energy H^0 and the energy of the mean must be accounted for entirely by the contribution of the fluctuations to the kinetic energy. The same conclusion also follows from equation (40). This result provides a new and definite interpretation to the notion set forth by Yankov *et al.* [9, 6, 7] and Pomeau [10] that the "entropy" of the NLS system is directly related to the amount of kinetic energy contained in the small–scale fluctuations, or radiation, and that the dynamical tendency of solutions to approach a ground state that minimizes energy at a conserved particle number is "thermodynamically advantageous."

The parameter $\lambda^{(n)}$ is determined as a Lagrange multiplier for the mean-field variational problem: minimize H_n subject to the constraint that $N_n = N^0$. The multiplier associated with any minimizer $\langle \phi^{(n)}\rangle = \langle u^{(n)}\rangle + i\langle v^{(n)}\rangle$ is unique, even though the minimizer itself may be nonunique. Indeed, the family of solutions $e^{i\theta}\langle \phi^{(n)}\rangle$ for arbitrary (constant) phase shifts θ exists for every such minimizer. Moreover, the ground state equation is a nonlinear eigenvalue equation, whose solutions can bifurcate. These bifurcations play the role of phase

transitions in the mean-field statistical theory. Some analysis of the solutions of the mean-field equation is provided in the Appendix.

Once a solution pair $(\langle\phi^{(n)}\rangle, \lambda^{(n)})$ is found, with $H_n(\langle\phi^{(n)}\rangle) = H_n^*$ being the minimum value of H_n allowed by the particle number constraint $N_n = N^0$, the "inverse temperature" $\beta^{(n)}$ is uniquely determined by

$$\beta^{(n)} = \frac{n}{H^0 - H_n^*}\,. \tag{43}$$

Indeed, this is a restatement of (41). Since H_n^* approaches a finite limit as $n \to \infty$, we find that $\beta^{(n)}$ scales linearly with the number of modes n. This asymptotic scaling of inverse temperature, $\beta^{(n)} \sim n\beta^*$, where β^* is finite in the continuum limit, is what distinguishes our mean-field theory from the statistical equilibrium theories discussed in Section 3. In those theories the inverse temperature is fixed and the mean energy is allowed to go to infinity in the continuum limit. Our rescaling of inverse temperature with n is necessary in order that the expectation of Hamiltonian with respect to $\rho^{(n)}$ remain finite in the limit as $n \to \infty$.

The implications of this scaling of $\beta^{(n)}$ are most evident in the the particle number and kinetic energy spectral densities. Substituting the expression (43) for $\beta^{(n)}$ into (38), we obtain the following formulas for the variances of the Fourier components u_k and v_k:

$$\mathrm{Var}\ (u_k) = \mathrm{Var}\ (v_k) = \frac{H^0 - H_n^*}{n\lambda_k}\,. \tag{44}$$

From these equilibrium expressions we obtain a sharp form of the vanishing of fluctuations hypothesis (21), which we adopted to derive the mean-field ensembles; namely,

$$\frac{1}{2}\sum_{k=1}^{n}[\mathrm{Var}(u_k) + \mathrm{Var}(v_k)] \;=\; \frac{H^0 - H_n^*}{n}\sum_{k=1}^{n}\frac{1}{\lambda_k} \;=\; O\left(\frac{1}{n}\right) \quad \text{as } n \to \infty\,, \tag{45}$$

using the fact that $\lambda_k^{-1} = O(k^{-2})$ in the estimate. Thus, we have the following prediction for the particle number spectral density

$$\frac{1}{2}\langle u_k^2 + v_k^2\rangle = \frac{1}{2}(\langle u_k\rangle^2 + \langle v_k\rangle^2) + \frac{H^0 - H_n^*}{n\lambda_k}\,. \tag{46}$$

Since the mean-field is a smooth solution of the ground state equation, its spectrum decays rapidly in k, and we have the approximation $\frac{1}{2}\langle u_k^2 + v_k^2\rangle \approx (H^0 - H_n^*)/(n\lambda_k)$ for $k >> 1/L$. In the same way, the kinetic energy spectral density is

$$\frac{1}{2}\lambda_k(\langle u_k^2 + v_k^2\rangle) = \frac{1}{2}\lambda_k(\langle u_k\rangle^2 + \langle v_k\rangle^2) + \frac{H^0 - H_n^*}{n}\,, \tag{47}$$

This expression shows that, as may be anticipated from a statistical equilibrium ensemble, the contribution to the kinetic energy from the fluctuations is equipartitioned among the n spectral modes. These predictions are tested against the results of direct numerical simulations in [8], where a good agreement with the mean-field theory is documented.

7 Continuum Limit of Random Fields

In this section we investigate the limits as $n \to \infty$ of the random fields $u^{(n)}$, $v^{(n)}$, and their gradients. The random fields $u^{(n)}$ and $v^{(n)}$ will be seen to converge in a uniform way to deterministic limits u^* and v^*, while the gradients u_x^* and v_x^* converge in a weaker sense to random fields with finite variance but no spatial coherence.

A simple form of the convergence of $u^{(n)}$ can be seen from the following easy estimate:

$$\begin{aligned}
\mathrm{Var}(u^{(n)}(x)) \;&=\; \sum_{k=1}^{n}\mathrm{Var}(u_k)\,(e_k(x))^2 \\
&=\; \frac{1}{\beta^{(n)}}\sum_{k=1}^{n}\frac{1}{\lambda_k}(e_k(x))^2 \\
&=\; O(n^{-1}),
\end{aligned} \tag{48}$$

and similarly for $v^{(n)}(x)$. Thus, the variance at each point goes to zero as $n \to \infty$.

A more detailed analysis is complicated by the fact that the mean fields $(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ are known to converge only on appropriately selected subsequences. This result is the content of Theorem A1 proved in the Appendix. It states that such subsequences of mean fields converge in the C^1-norm to ground states (u^*, v^*) that solve the continuum equation (2); these ground states minimize the Hamiltonian H over all $(u, v) \in H_0^1(\Omega)$ that satisfy the particle number constraint $N(u, v) = N^0$. The possible nonuniqueness of solutions to the ground state equation necessitates the introduction of subsequences, since in general there may exist a *set* of limits (u^*, v^*) for a specified constraint value N^0. Accordingly, we shall assume throughout this section that n goes to infinity along a subsequence for which the limiting ground state (u^*, v^*) exists, without relabeling n.

The detailed properties of the continuum limit are expressed in the following theorem.

Theorem 2. The (sub)sequence of random fields $(u^{(n)}, v^{(n)})$ whose distribution is the mean-field ensemble with density $\rho^{(n)}$ converges in distribution to a non–random field (u^*, v^*) which minimizes the Hamiltonian H given the particle number constraint $N = N^0$. In addition, the gradients $(u_x^{(n)}, v_x^{(n)})$ converge in the sense of finite-dimensional distributions to a Gaussian random field $(U'(x), V'(x))$ on $\Omega = [0, L]$ with the following properties: U' and V' are statistically independent, the mean of (U', V') is (u_x^*, v_x^*), and U' and V' each have the covariance function

$$\Gamma^*(x,y) = \frac{H^0 - H^*}{L} \begin{cases} 0, & \text{if } x \neq y \\ 1, & \text{if } x = y \text{ and } 0 < x < L \\ 2, & \text{if } x = y \text{ and } x = 0, L \end{cases} \tag{49}$$

where H^* is the minimum value of H given $N = N^0$.

Proof. Recall that for each n, the mean $(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ under $\rho^{(n)}$ is a solution to (39) which minimizes H_n over $X_n \times X_n$ subject to the constraint $N_n = N^0$. By Theorem A1, the (sub)sequence $(\langle u^{(n)}\rangle, \langle v^{(n)}\rangle)$ converges in $C^1(\Omega)$ to (u^*, v^*), where (u^*, v^*) is a solution of (9) that minimizes H subject to the constraint $N(u, v) = N^0$.

Recalling (38) and the independence of the Fourier coefficients u_k, we may use the following representation of $u^{(n)}$:

$$\begin{aligned} u^{(n)} = \sum_{k=1}^{n} u_k e_k &= \sum_{k=1}^{n} \left[\langle u_k \rangle + \frac{1}{\sqrt{\lambda_k \beta^{(n)}}} Z_k \right] e_k \\ &= \langle u^{(n)} \rangle + \frac{1}{\sqrt{\beta^{(n)}}} \sum_{k=1}^{n} \frac{L}{\pi k} Z_k e_k, \end{aligned} \tag{50}$$

where $Z_1, Z_2, \ldots$ are independent Gaussian random variables with mean 0 and variance 1. The sum in the last line is the n^{th} partial sum of the Karhunen-Loève expansion of the Brownian bridge on Ω, which is known to converge to its limit $B(x), x \in \Omega$, uniformly with probability one as $n \to \infty$ ([35], Theorem 3.3.3). That is, for each function $h : C(\Omega) \to R$ which is continuous in the supremum norm, we have, with probability one,

$$h\left(\sum_{k=1}^{n} \frac{L}{\pi k} Z_k e_k \right) \to h(B) \;\text{ as }\; n \to \infty. \tag{51}$$

It is clear from this demonstration that for such h we have

$$h(u^{(n)}) \to h(u^*) \;\text{ as }\; n \to \infty \tag{52}$$

with probability one, since along this (sub)sequence the means converge uniformly and $\beta^{(n)}$ goes to infinity. Restricting to bounded h and taking averages, we have

$$\langle h(u^{(n)}) \rangle \to h(u^*) \;\text{ as }\; n \to \infty, \tag{53}$$

which is the definition of convergence in distribution of the random fields $u^{(n)}$ to u^*. The argument that $v^{(n)}$ converges to v^* is identical.

The analysis of the gradient fields $u_x^{(n)}$ and $v_x^{(n)}$ is simpler. First note that they are independent and differ only in their means, and consequently it is sufficient to consider $u_x^{(n)}$, say. From the expression

$$u_x^{(n)}(x) = \sum_{k=1}^{n} u_k (e_k)_x(x) \,, \tag{54}$$

it is evident that $u_x^{(n)}$ is a continuous Gaussian random field whose mean $\langle u_x^{(n)} \rangle$ converges uniformly to u_x^* as $n \to \infty$. Its covariance is then calculated to be

$$\begin{aligned} \Gamma_n(x,y) &= \frac{1}{\beta^{(n)}} \sum_{k=1}^{n} \frac{1}{\lambda_k} (e_k)_x(x)(e_k)_x(y) \\ &= \frac{1}{2L\beta^{(n)}} \left[D_n(\pi(x+y)/L) + D_n(\pi(x-y)/L) - 2 \right] \,, \end{aligned} \tag{55}$$

where

$$D_n(\theta) = \begin{cases} \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}, & \text{for } \theta \neq 0, \pm 2\pi, \pm 4\pi, \ldots \\ 2n+1, & \text{for } \theta = 0, \pm 2\pi, \pm 4\pi, \ldots \end{cases}$$

is the Dirichlet kernel [32]. The asymptotic behavior $\beta^{(n)} \sim n/(H^0 - H^*)$ implied by (43) combined with the properties of $D_n(\theta)$ then produce the desired result that, as $n \to \infty$, the covariance function $\Gamma_n(x,y)$ converges pointwise to the function $\Gamma^*(x,y)$ given in the statement of the theorem.

Now let U' be the Gaussian random field with mean u_x^* and covariance Γ^*. For convergence of finite-dimensional distributions, it remains to show that for each m and any $x_1, \ldots, x_m$ in Ω, the joint distribution on R^m of the vector $(u_x^{(n)}(x_1), \ldots, u_x^{(n)}(x_m))$ converges weakly to the joint distribution of $(U'(x_1), \ldots, U'(x_m))$ as $n \to \infty$. But this follows by the Lèvy Continuity Theorem [34], since the characteristic function of the former vector converges to the characteristic function of the latter vector, simply because the means and covariances converge. The argument that $v_x^{(n)}$ converges to V' is similar, as is the extension to the convergence of the pair $(u_x^{(n)}, v_x^{(n)})$.

The results of Theorem 2 may be paraphrased as follows: The mean-field distributions $\rho^{(n)}$ converge weakly as measures on $C(\Omega)$ (along a subsequence) to a Dirac mass concentrated at a minimizer of the Hamiltonian subject to the particle number constraint. In this limit, the variance of the field $\phi^{(n)}$ vanishes while the variance of its gradient $\phi_x^{(n)}$ is uniform over the interval Ω, apart from the endpoints. This behavior is in good qualitative agreement with the long–time behavior observed in numerical simulations [8].

Nonetheless, the continuum limit is degenerate in two ways. First, the limiting field ϕ^* is deterministic. Second, the limiting gradient field, whose mean is ϕ_x^* and covariance function is Γ^*, is so rough that it does not even take values in the space of measurable functions (that is, it does not have a measurable version) [33]. These unusual properties of the limiting distribution are a direct consequence of the scaling of the inverse temperature $\beta^{(n)}$ with n, which is required to maintain finite mean energy in the continuum limit.

8 The Concentration Property

We shall establish in this section the important result that the mean-field ensembles $\rho^{(n)}$, which solve the maximum entropy principle (MEP), become equivalent in a certain sense to the microcanonical ensemble (19) in the continuum limit $n \to \infty$. Specifically, we shall prove the following *concentration property*:

Theorem 3. Let $\rho^{(n)}, n = 1, 2, \cdots$, be a sequence of solutions of the constrained variational principle (MEP). Then

$$\lim_{n\to\infty} \langle N_n \rangle = N^0 \,, \quad \lim_{n\to\infty} \mathrm{Var}(N_n) = 0 \,, \tag{56}$$

and

$$\lim_{n\to\infty} \langle H_n \rangle = H^0 \,, \quad \lim_{n\to\infty} \mathrm{Var}(H_n) = 0 \,. \tag{57}$$

This theorem has the interpretation that the mean-field ensembles $\rho^{(n)}$ concentrate on the microcanonical constraint manifold $N_n = N^0, H_n = H^0$ in the continuum limit. As we have emphasized above, it is a well-accepted axiom of statistical mechanics that the microcanonical ensemble constitutes the appropriate statistical equilibrium description for an isolated ergodic system [11, 30]. Thus, this concentration property provides an important theoretical justification for our mean-field theory.

In the proof of Theorem 3, we will make use of the following elementary facts concerning Gaussian random variables: If W is a Gaussian random variable with mean μ and variance σ^2, then

$$\langle (W-\mu)^4 \rangle = 3\sigma^4 \,, \quad \langle W^4 \rangle - \langle W^2 \rangle^2 = 2\sigma^4 + 4\sigma^2\mu^2 \,. \tag{58}$$

We will also repeatedly make use of the result that $1/\beta^{(n)} = O(n^{-1})$, which follows from equation (43).

Proof of Theorem 3: The first conclusion in (56) is easy to establish. Indeed, using (24) and the calculation (45), we have

$$\begin{aligned} \langle N_n \rangle &= \frac{1}{2}\sum_{k=1}^{n} (\langle u_k \rangle^2 + \langle v_k \rangle^2) + \sum_{k=1}^{n} [\mathrm{Var}(u_k) + \mathrm{Var}(v_k)] \\ &= N^0 + O(n^{-1}) \,, \text{ as } n \to \infty \,. \end{aligned}$$

To verify the second assertion in (56), we calculate using (18)

$$\begin{aligned} \mathrm{Var}(N_n) &= \frac{1}{4} \left\langle \left(\sum_{k=1}^{n} \left[(u_k^2 - \langle u_k^2 \rangle) + (v_k^2 - \langle v_k^2 \rangle) \right] \right)^2 \right\rangle \\ &= \frac{1}{4} \sum_{k=1}^{n} \left[(\langle u_k^4 \rangle - \langle u_k^2 \rangle^2) + (\langle v_k^4 \rangle - \langle v_k^2 \rangle^2) \right] \,, \end{aligned} \tag{59}$$

where, to obtain the second equality in this calculation, we have made use of the mutual statistical independence of the random variables $u_1, \ldots, u_n, v_1, \ldots, v_n$. But, as u_k and v_k are Gaussian with means $\langle u_k \rangle$ and $\langle v_k \rangle$ and variances given by (44), we have from (58) that

$$\langle u_k^4 \rangle - \langle u_k^2 \rangle^2 = \frac{2}{(\beta^{(n)}\lambda_k)^2} + \frac{4\langle u_k \rangle^2}{\beta^{(n)}\lambda_k} \,, \quad \langle v_k^4 \rangle - \langle v_k^2 \rangle^2 = \frac{2}{(\beta^{(n)}\lambda_k)^2} + \frac{4\langle v_k \rangle^2}{\beta^{(n)}\lambda_k} \,. \tag{60}$$

Upon substituting (60) into (59), we obtain

$$\begin{aligned} \mathrm{Var}(N_n) &= \frac{1}{(\beta^{(n)})^2} \sum_{k=1}^{n} \frac{1}{\lambda_k^2} + \frac{1}{\beta^{(n)}} \sum_{k=1}^{n} \left[\frac{\langle u_k \rangle^2 + \langle v_k \rangle^2}{\lambda_k} \right] \\ &\to 0 \quad \text{as} \quad n \to \infty \,, \end{aligned}$$

since $\sum_{k=1}^{n} \lambda_k^{-2}$ converges as $n \to \infty$, and $\sum_{k=1}^{n} \lambda_k^{-1}(\langle u_k \rangle^2 + \langle v_k \rangle^2) \le 2\lambda_1^{-1} N^0$ for all n.

To establish the first part of (57), observe that from (25), there holds

$$\left| \langle H_n \rangle - H^0 \right| = \left| \langle \Theta_n(u^{(n)}, v^{(n)}) \rangle - \Theta_n(\langle u^{(n)} \rangle, \langle v^{(n)} \rangle) \right| \,, \tag{61}$$

where the potential energy Θ_n is defined by (17). Mimicking the calculations following equation (26) and preceding equation (29), we find that

$$\left|\langle\Theta_n(u^{(n)},v^{(n)})\rangle - \Theta_n(\langle u^{(n)}\rangle,\langle v^{(n)}\rangle)\right| \le C\sum_{k=1}^{n}\left[\mathrm{Var}(u_k)+\mathrm{Var}(v_k)\right] , \tag{62}$$

where C is a constant independent of n. Thus, according to (61), (62) and (45), there holds

$$\left|\langle H_n\rangle - H^0\right| = O(n^{-1}) , \ \text{ as } \ n\to\infty .$$

Next, we have that

$$\mathrm{Var}(H_n) = \mathrm{Var}(K_n+\Theta_n) \le 2\ \mathrm{Var}(K_n) + 2\ \mathrm{Var}(\Theta_n) , \tag{63}$$

where K_n is the kinetic energy defined as in (16). Using the statistical independence properties noted above, we calculate

$$\begin{aligned}\mathrm{Var}(K_n) &= \frac{1}{4}\left\langle\left(\sum_{k=1}^{n}\lambda_k\left[(u_k^2-\langle u_k^2\rangle)+(v_k^2-\langle v_k^2\rangle)\right]\right)^2\right\rangle \\ &= \frac{1}{4}\sum_{k=1}^{n}\lambda_k^2\left[(\langle u_k^4\rangle-\langle u_k^2\rangle^2)+(\langle v_k^4\rangle-\langle v_k^2\rangle^2)\right] ,\end{aligned}$$

and using (60), we arrive at

$$\mathrm{Var}(K_n) = \frac{n}{(\beta^{(n)})^2} + \frac{1}{\beta^{(n)}}\sum_{k=1}^{n}\lambda_k(\langle u_k\rangle^2+\langle v_k\rangle^2) .$$

This last expression vanishes as $n\to\infty$, because the sum $\sum_{k=1}^{n}\lambda_k(\langle u_k\rangle^2+\langle v_k\rangle^2)$, which represents twice the kinetic energy of the mean, is bounded independently of n.

Finally, we show that the variance of the potential energy Θ_n tends to 0 in the continuum limit. Using the definition of Θ_n and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\mathrm{Var}(\Theta_n) &= \frac{1}{4}\left\langle\left(\int_\Omega\left[F((u^{(n)})^2+(v^{(n)})^2) - \left\langle F((u^{(n)})^2+(v^{(n)})^2)\right\rangle\right]\, dx\right)^2\right\rangle \\ &\le \frac{L}{4}\int_\Omega \mathrm{Var}\left(F((u^{(n)})^2+(v^{(n)})^2)\right)\, dx .\end{aligned} \tag{64}$$

Now the potential F may be expanded about the mean $(\langle u^{(n)}\rangle,\langle v^{(n)}\rangle)$ as

$$\begin{aligned}F((u^{(n)})^2+(v^{(n)})^2) &= F(\langle u^{(n)}\rangle^2+\langle v^{(n)}\rangle^2) \\ &+ 2f(\langle u^{(n)}\rangle^2+\langle v^{(n)}\rangle^2)\left[\langle u^{(n)}\rangle(u^{(n)}-\langle u^{(n)}\rangle)+\langle v^{(n)}\rangle(v^{(n)}-\langle v^{(n)}\rangle)\right] \\ &+ \frac{1}{2}\begin{pmatrix} (u^{(n)}-\langle u^{(n)}\rangle) \\ (v^{(n)}-\langle v^{(n)}\rangle)\end{pmatrix}^T J(\tilde u^{(n)},\tilde v^{(n)})\begin{pmatrix} (u^{(n)}-\langle u^{(n)}\rangle) \\ (v^{(n)}-\langle v^{(n)}\rangle)\end{pmatrix} ,\end{aligned}$$

where the matrix J is defined by (28) and $(\tilde u^{(n)},\tilde v^{(n)})$ lies between $(u^{(n)},v^{(n)})$ and $(\langle u^{(n)}\rangle,\langle v^{(n)}\rangle)$. Using the boundedness condition (5) on f and the statistical independence of $u^{(n)}$ and $v^{(n)}$, we obtain after a straightforward analysis, that pointwise over Ω

$$\begin{aligned}\mathrm{Var}\left(F((u^{(n)})^2+(v^{(n)})^2)\right) &\leq C\left[\langle u^{(n)}\rangle^2\,\mathrm{Var}(u^{(n)})+\langle v^{(n)}\rangle^2\,\mathrm{Var}(v^{(n)})\right] \\ &+ C\left[\langle (u^{(n)}-\langle u^{(n)}\rangle)^4\rangle+\langle (v^{(n)}-\langle v^{(n)}\rangle)^4\rangle\right]. \end{aligned} \tag{65}$$

Here, C is a constant that is independent of both n and $x \in \Omega$. As we have demonstrated in (48),

$$\mathrm{Var}(u^{(n)}(x)) = \mathrm{Var}(v^{(n)}(x)) = \frac{1}{\beta^{(n)}}\sum_{k=1}^{n}\frac{e_k^2(x)}{\lambda_k}\,, \tag{66}$$

and since $u^{(n)}(x)$ and $v^{(n)}(x)$ are Gaussian variables, it follows from (66) and (58) that

$$\left\langle (u^{(n)}(x)-\langle u^{(n)}(x)\rangle)^4\right\rangle = \left\langle (v^{(n)}(x)-\langle v^{(n)}(x)\rangle)^4\right\rangle = \frac{3}{(\beta^{(n)})^2}\left(\sum_{k=1}^{n}\frac{e_k^2(x)}{\lambda_k}\right)^2 . \tag{67}$$

Thus, from (64)–(67), the fact that the eigenfunctions e_k are uniformly bounded over Ω independently of k, and the fact that $\int_\Omega(\langle u^{(n)}\rangle^2+\langle v^{(n)}\rangle^2)\,dx = 2N^0$ for all n, we have

$$\mathrm{Var}(\Theta_n) = O(n^{-1})\,, \text{ as } n\to\infty.$$

This completes the proof of Theorem 3.

9 Extension to Unbounded Nonlinearities

In this section, we briefly indicate how our mean-field statistical theory, can be extended to a class of unbounded nonlinearities, which includes the focusing power law nonlinearities $f(|\psi|^2)=|\psi|^s, 0<s<4$. As illustrated in [8, 6, 7], numerical simulations for such power law nonlinearities exhibit the same phenomena as seen for bounded nonlinearities such as $f(|\psi|^2)=|\psi|^2/(1+|\psi|^2)$, for which the results of numerical simulations are displayed in Figures (1)-(2). That is, the field ψ approaches a long-time state consisting of a coherent soliton structure coupled with radiation or fluctuations of very small amplitude. As in the case of the bounded nonlinearities, the gradient ψ_x exhibits fluctuations of nonnegligible amplitude for all time. In fact, we expect, that this general behavior occurs as long as the nonlinearity f is such that the NLS equation (1) is nonintegrable and free of collapse (i.e., such that finite time singularity does not occur). Hence, we would like to apply our statistical theory to NLS with such nonlinearities, as well.

Let us recall that to motivate the mean-field constraints (24) and (25) for nonlinearities f satisfying the boundedness condition (5), we invoked the vanishing of fluctuations hypothesis (21). However, for nonlinearities such that the potential $F(|\psi|^2)$ grows more rapidly than $C|\psi|^2$ as $|\psi|\to\infty$ (as it does for the power law nonlinearities), the hypothesis (21) is not sufficient to guarantee *a priori* that $\langle\Theta_n(u^{(n)},v^{(n)})\rangle$ converges to $\Theta_n(\langle u^{(n)}\rangle,\langle v^{(n)}\rangle)$ as $n\to\infty$. Thus the mean-field Hamiltonian constraint (25) can not be derived from the vanishing of fluctuations hypothesis (21) alone for such f. Note that by making a stronger vanishing of fluctuations hypothesis, we could have weakened the assumptions on f and arrived at the same mean-field constraints (24)-(25). In any case, we could simply impose these mean-field constraints and investigate the resulting maximum entropy ensembles $\rho^{(n)}$. If it can be shown that these ensembles exist and satisfy the concentration property expressed in Theorem 3, then we will consider this approach to be justified *a posteriori*.

It is clear from Theorem 1 and Theorem 2 that, in order for our statistical theory to be well-defined, it is necessary that there exist minimizers (in $H_0^1(\Omega)$, say) of the Hamiltonian H given the particle number constraint $N=N^0$. This places restrictions on the class of nonlinearities that we can consider. However for nonlinearities f such that the potential F satisfies

$$\text{There exists } C>0 \text{ such that } F(|\psi|^2)\leq C(|\psi|^2+|\psi|^q)\,, \text{ for some } 2\leq q<6\,, \tag{68}$$

it may be shown that such minimizers exist for any N^0 [36]. The condition (68) is also crucial in establishing the well-posedness of the NLS equation (1) as an initial value problem in the Sobolev space $H_0^1(\Omega)$ [28]. Clearly, the focusing power law nonlinearities $f(|\psi|^2) = |\psi|^s$ satisfy this condition when $0 < s < 4$.

Thus, assuming that F satisfies the growth condition (68), we may investigate the ensembles $\rho^{(n)}$ which maximize entropy subject to the mean-field constraints (24)-(25). It is easy to see that the analysis of Section 6 goes through without change. Thus, under the the maximum entropy ensemble $\rho^{(n)}$, the Fourier coefficients $u_1, \cdots u_n, v_1, \cdots v_n$ are mutually independent Gaussian variables, the mean-field minimizes the Hamiltonian given the particle number constraint $N_n = N^0$, and the variances of the u_k and v_k are given by (44). It may be shown that the continuum limit of Section 7 and the important concentration property of Section 8 also hold, if we impose, in addition to (68), a lower bound on the potential F of the form

$$\text{There exists } c \geq 0 \text{ such that } -c|\psi|^r \leq F(|\psi|^2)\,, \text{ for some } r \geq 0\,. \tag{69}$$

That is, Theorems 2 and 3 still hold for F satisfying (68) and (69), but the proofs are complicated by the unboundedness of f. Since our goal in this paper has been to emphasize conceptual issues rather than technical details, we shall not present the proofs here. However, we refer the interested reader to [37], where some of the analysis has been carried out for nonlinearities satisfying (68) and (69).

Acknowledgments

The research of R. Jordan has been supported in part by an NSF Postdoctoral Research Fellowship and also by DOE through grants to the Center for Nonlinear Studies at Los Alamos National Laboratory. The research of B. Turkington is partially supported by the NSF through the grant NSF-9600060. R. Jordan thanks Christophe Josserand and Pieter Swart for valuable discussions and suggestions.

Appendix

In this Appendix, we provide an analysis of the variational principle that defines the mean field corresponding to the maximum entropy ensemble $\rho^{(n)}$. In particular, we prove that for (smooth) nonlinearities f satisfying the condition (5), solutions of this variational principle exist, and we investigate the convergence properties of these mean fields as $n \to \infty$.

First, we consider the variational principle

$$H(\psi) \to \min\,, \text{ subject to } \psi \in A^0\,, \tag{70}$$

where the admissible set A^0 is defined as

$$A^0 = \left\{\psi : \Omega \to \mathbf{C} \mid \psi \in H_0^1(\Omega)\,,\ N(\psi) = N^0\right\}\,. \tag{71}$$

The Hamiltonian H and the particle number N are defined by equations (3) and (4), respectively, and, as in the main text, $\Omega = [0, L]$. Recall also that the potential F is defined as $F(a) = \int_0^a f(a')\,da'$.

Lemma A1. There exists a solution $\psi \in H_0^1(\Omega)$ of the variational problem (70).

Proof: As f satisfies the boundedness condition (5), there exists a constant $K > 0$ such that $|F(|\psi|^2)| \leq K|\psi|^2$. Thus, for all $\psi \in A^0$, there holds

$$\begin{aligned} H(\psi) &= \frac{1}{2}\int |\psi_x|^2\,dx - \frac{1}{2}\int F(|\psi|^2)\,dx \\ &\geq \frac{1}{2}\int |\psi_x|^2\,dx - \frac{K}{2}N^0\,. \end{aligned} \tag{72}$$

Here, and in the remainder of this Appendix, all integrals are over the domain Ω. It follows from (72) that H is bounded below on A^0, so that we may choose a minimizing sequence $(\psi_n)_{n\in\mathbf{N}} \subset A^0$ for H. The inequality (72) also implies that the sequence ψ_n is uniformly bounded in $H^1_0(\Omega)$. Hence, by the Sobolev Imbedding Theorem [38], there exists a subsequence, still denoted by ψ_n, and a function ψ such that ψ_n converges weakly to ψ in $H^1_0(\Omega)$ and ψ_n converges in $C(\Omega)$ to ψ. Because of the convergence in $C(\Omega)$, we have that $\int F(|\psi|^2)\,dx = \lim_{n\to\infty} \int F(|\psi_n|^2)\,dx$. Also, since ψ_n converges weakly to ψ in $H^1_0(\Omega)$, there holds $\int |\psi_x|^2\,dx \le \liminf_{n\to\infty} \int |(\psi_n)_x|^2\,dx$. Putting these two results together, we obtain that

$$H(\psi) \le \liminf_{n\to\infty} H(\psi^{(n)}) = \inf_{\phi\in A^0} H(\phi)\,. \tag{73}$$

But, as ψ_n converges to ψ in $C(\Omega)$, it follows that $N(\psi) = N^0$, so that $\psi \in A^0$. We conclude from (73), therefore, that ψ is a solution of the variational problem (70).

Lemma A2. Any solution ψ of the variational problem (70) is a solution of the ground state equation (2) for some real λ.

Proof: The conclusion of this lemma follows from an application of the Lagrange multiplier rule for constrained optimization. The parameter λ in the ground state equation (2) is the Lagrange multiplier which enforces the constraint $N(\psi) = N^0$.

We have demonstrated in Section (6) that the mean fields $\langle\phi^{(n)}\rangle = \langle u^{(n)}\rangle + i\langle v^{(n)}\rangle$ corresponding to the maximum entropy ensembles $\rho^{(n)}$ are solutions of the following variational problem

$$H(\psi) \to \min\,, \;\; \text{subject to } \psi \in A^0_n\,, \tag{74}$$

where the admissible set A^0_n is defined as

$$A^0_n = \left\{\psi \in \Xi_n \mid N(\psi) = N^0\right\}\,. \tag{75}$$

Here, Ξ_n is the set of all complex fields ψ on Ω having the form $\psi = u + iv$, where the real fields u and v are elements of the n-dimensional space X_n spanned by the first n eigenfunctions $e_1, \cdots, e_n$ of the operator $-d^2/dx^2$ on Ω with homogeneous Dirichlet boundary conditions. Ξ_n may be thought of as a closed subspace of the Sobolev space $H^1_0(\Omega)$. The sequence of variational problems (74) corresponds to a Ritz–Galerkin scheme for approximating solutions of the variational principle (70). The proof of the next lemma follows from arguments analogous to those that were used to prove Lemmas A1 and A2. The conclusion about the smoothness of solutions of (74) is obvious, because any such solution is necessarily a finite linear combination of the basis functions $e_k(x) = \sqrt{2/L}\sin(k\pi x/L)$.

Lemma A3. For each n, there exist solutions of the variational problem (74). Any such solution $\psi^{(n)}$ is a solution of the differential equation

$$\psi^{(n)}_{xx} + P^{(n)}\left(f(|\psi^{(n)}|^2)\psi^{(n)}\right) + \lambda^{(n)}\psi^{(n)} = 0\,, \tag{76}$$

where $P^{(n)}$ is the projection from $H^1_0(\Omega)$ onto Ξ_n, and $\lambda^{(n)}$ is a Lagrange multiplier which enforces the constraint $N(\psi^{(n)}) = N^0$. In addition, any solution $\psi^{(n)}$ of (74) is in $C^\infty(\Omega)$.

We wish to investigate the convergence of solutions of (74) to those of (70). It is not our objective to obtain the strongest possible convergence results. For our purposes, it is sufficient to prove that any sequence of solutions of (74) has a subsequence which converges in $C^1(\Omega)$ to a solution of (70). Such a result is all that is needed for the analysis of the continuum limit of the mean field ensembles $\rho^{(n)}$ in Section (7). In general, we do not know that solutions of the variational principle are unique, so that subsequences can not be avoided. We shall now prove the following theorem.

Theorem A1. If $\psi^{(n)}, n = 1, 2, \cdots$ is a sequence of solutions of the variational problem (74), then there exists a subsequence $\psi^{(n')}$ of $\psi^{(n)}$ which converges in $C^1(\Omega)$ as $n' \to \infty$ to a solution of the variational problem (70).

Proof: Let $\psi^{(n)}$ be a sequence of solutions of (74), and let $H_n^* = H(\psi^{(n)})$. Because $A_n^0 \subset A_{n+1}^0, n = 1, 2, \cdots$, and $A^0 = \overline{\bigcup_{n=1}^{\infty} A_n^0}$, we have that $H_n^* \geq H_{n+1}^*$, and $H_n^* \searrow H^*$ as $n \to \infty$, where H^* is the minimum value of H over A^0. Thus, $\psi^{(n)}$ is a minimizing sequence for the variational problem (70), and so following the proof of Lemma A1, there exists a subsequence $\psi^{(n')}$ and a solution ψ of the variational problem (70) such that $\psi^{(n')}$ converges weakly in $H_0^1(\Omega)$ to ψ and $\psi^{(n')}$ converges in $C(\Omega)$ to ψ.

Now, as $\psi^{(n')}$ is a solution of the variational problem (74), it satisfies equation (76) for some Lagrange multiplier $\lambda^{(n')}$. We shall now show that the sequence $\lambda^{(n')}, n' \to \infty$ is bounded independently of n'. For this purpose, we multiply equation (76) for $\psi^{(n')}$ by the complex conjugate $\overline{\psi^{(n')}}$ and integrate over Ω to obtain

$$-\int |(\psi^{(n')})_x|^2\, dx + \int P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')})\overline{\psi^{(n')}}\, dx + \lambda^{(n')} \int |\psi^{(n')}|^2\, dx = 0\,. \tag{77}$$

It follows immediately from (77) and the fact that $\int |\psi^{(n')}|^2\, dx = 2N^0$ that

$$|\lambda^{(n')}| \leq \frac{1}{2N^0}\left[\int |(\psi^{(n')})_x|^2\, dx + \left|\int P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')})\overline{\psi^{(n')}}\, dx\right|\right]. \tag{78}$$

Because $\psi^{(n')}$ converges weakly to ψ in $H_0^1(\Omega)$, the first integral on the right hand side of (78) is bounded independently of n'. The second integral on the right hand side of (78) may be estimated as follows:

$$\begin{aligned}
\left|\int P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')})\overline{\psi^{(n')}}\, dx\right| &\leq \left(\int |\psi^{(n')}|^2\, dx\right)^{\frac{1}{2}} \left(\int |P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')})|^2\, dx\right)^{\frac{1}{2}} \\
&\leq \left(\int |\psi^{(n')}|^2\, dx\right)^{\frac{1}{2}} \left(\int (f(|\psi^{(n')}|^2))^2 |\psi^{(n')})|^2\, dx\right)^{\frac{1}{2}} \\
&\leq 2N^0 \sup_{x\in\Omega} |f(|\psi^{(n')}|^2)|\,.
\end{aligned}$$

To obtain the first line of this display, we have used the Cauchy–Schwarz inequality. The second line follows from the fact that $\int |P^{(n)}\phi|^2\, dx \leq \int |\phi|^2\, dx$ for all $\phi \in H_0^1(\Omega), n \in \mathbf{N}$, and to obtain the third line, we have once again used the identity $\int |\psi^{(n')}|^2\, dx = 2N^0$. Now, owing to (5), $\sup_{x\in\Omega} |f(|\psi^{(n')}|^2)|$ is bounded independently of n', so the preceding calculation demonstrates that the second integral on the right hand side of (78) can also be bounded independently of n'. This implies that $\lambda^{(n')}$ is uniformly bounded in n'.

The strategy now is to use the uniform boundedness of the eigenvalues $\lambda^{(n')}$ to establish that the subsequence $\psi^{(n')}$ is uniformly bounded in the Sobolev space $H_0^2(\Omega)$. From this result and the Sobolev Imbedding Theorem [38], we may conclude that the $\psi^{(n')}$ converges in $C^1(\Omega)$ to the function ψ as above, which is a solution of the variational problem (70). This is the desired conclusion.

To prove that $\psi^{(n')}$ is, in fact, uniformly bounded in $H_0^2(\Omega)$, we multiply the equation (76) for $\psi^{(n')}$ by the complex conjugate of $(\psi^{(n')})_{xx}$ and integrate over Ω. This yields, after integrating by parts and using the homogeneous Dirichlet boundary conditions,

$$\int |(\psi^{(n')})_{xx}|^2\, dx = \lambda^{(n')} \int |(\psi^{(n')})_x|^2\, dx + \int \overline{\psi^{(n')}}_x (P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')}))_x\, dx\,. \tag{79}$$

Using the Cauchy–Schwarz inequality, and the inequality $\int |(P^n(\phi))_x|^2\, dx \leq \int |\phi_x|^2\, dx$, which is valid for all $n \in \mathbf{N}$ and all $\phi \in H_0^1(\Omega)$, we find that

$$\left|\int \overline{\psi^{(n')}}_x (P^{(n')}(f(|\psi^{(n')}|^2)\psi^{(n')}))_x\, dx\right| \leq \left(\int |(\psi^{(n')})_x|^2\, dx\right)^{\frac{1}{2}} \left(\int |(f(|\psi^{(n')}|^2)\psi^{(n')})_x|^2\, dx\right)^{\frac{1}{2}}. \tag{80}$$

But, a straightforward calculation yields that, pointwise on Ω,

$$\begin{aligned}
|(f(|\psi^{(n')}|^2)\psi^{(n')})_x|^2 &\leq 2(f(|\psi^{(n')}|^2))^2|(\psi^{(n')})_x|^2 + 8(f'(|\psi^{(n')}|^2))^2|\psi^{(n')}|^4|(\psi^{(n')})_x|^2 \\
&\leq C|(\psi^{(n')})_x|^2\,,
\end{aligned} \tag{81}$$

for a constant C independent of n'. The second line of this calculation follows from the boundedness condition (5) on f. Taken together, (79)-(81) imply that

$$\int |(\psi^{(n')})_{xx}|^2\, dx \leq (C + |\lambda^{(n')}|) \int |(\psi^{(n')})_x|^2\, dx\,. \tag{82}$$

As we have established above, $\lambda^{(n')}$ and $\int |(\psi^{(n')})_x|^2\, dx$ are bounded independently of n', so that from (82), we conclude that the subsequence $\psi^{(n')}$ is uniformly bounded in $H_0^2(\Omega)$. This concludes the proof of Theorem A1.

References

[1] A. Hasegawa. Self-organization processes in continuous media. *Adv. Phys.*, 34:1, 1985.

[2] J. C. McWilliams. The emergence of isolated vortices in turbulent flow. *J. Fluid Mech.*, 146:21, 1984.

[3] E. Segre and S. Kida. Late states of incompressible 2D decaying vorticity fields. *Fluid Dyn. Res.*, 23:89, 1998.

[4] D. Biskamp and H. Welter. Dynamics of decaying two–dimensional magnetohydrodynamic turbulence. *Phys. Fluids B*, 1:1964, 1989.

[5] R. Kinney, J. C. McWilliams, and T. Tajima. Coherent structures and turbulent cascades in two–dimensional incompressible magnetohydrodynamic turbulence. *Phys. Plasmas*, 2:3623, 1995.

[6] S. Dyachenko, V. E. Zakharov, A. N. Pushkarev, V. F. Shvets, and V. V. Yan'kov. Soliton turbulence in nonintegrable wave systems. *Soviet Phys. JETP*, 69:1144, 1989.

[7] V. E. Zakharov, A. N. Pushkarev, V. F. Shvets, and V. V. Yan'kov. Soliton turbulence. *JETP Lett.*, 48:83, 1988.

[8] R. Jordan and C. Josserand. Soliton turbulence and coherent structures in the nonlinear Schrödinger equation. 1999. In preparation.

[9] S. F. Krylov and V. V. Yan'kov. The role of solitons in strong turbulence. *Soviet Phys. JETP*, 52:41, 1980.

[10] Y. Pomeau. Asymptotic time behavior of solutions of nonlinear classical field equations. *Nonlinearity*, 5:707, 1992.

[11] R. Balescu. *Equilibrium and Nonequilibrium Statistical Mechanics.* Wiley, New York, 1975.

[12] J. Miller, P. B. Weichman, and M. C. Cross. Statistical mechanics, Euler equation, and Jupiter's red spot. *Phys. Rev. A*, 45:2328, 1992.

[13] R. Jordan and B. Turkington. Ideal magnetofluid turbulence in two dimensions. *J. Stat. Phys.*, 87:661, 1997.

[14] M. Ablowitz and H. Segur. On the evolution of packets of water waves. *J. Fluid Mech.*, 92:691, 1979.

[15] H. L. Pecseli. Solitons and weakly nonlinear waves in plasmas. *IEEE Trans. Plasma Sci*, 13:53, 1985.

[16] A. Hasegawa and Y. Kodama. *Solitons in Optical Communications.* Oxford University Press, New York, 1995.

[17] H. Hasimoto. A soliton on a vortex filament. *J. Fluid Mech.*, 139:67, 1972.

[18] A. Majda. Simplified asymptotic equations for slender vortex filaments. *Proc. Symp. Appl. Math.*, 54:237, 1998.

[19] Y. Li and D. McLaughlin. Morse and Melnikov functions for NLS pdes. *Commun. Math. Phys.*, 162:175, 1994.

[20] V. E. Zakharov and A. B. Shabat. Exact theory of two–dimensional self–focusing and one–dimensional self–modulation of waves in nonlinear media. *Soviet Phys. JETP*, 34:62, 1972.

[21] D. Anderson and M. Bonnedal. Variational approach to nonlinear self focusing of Gaussian laser beams. *Phys. Fluids*, 22:105, 1979.

[22] D. Mihalache, R. G. Nazmitdinov, and V. K. Fedyanin. Nonlinear optical waves in layered structures. *Soviet J. Part. Nucl.*, 20:86, 1989.

[23] C. E. Max. Strong self–focusing due to ponderomotive forces in plasmas. *Phys. Fluids*, 19:74, 1976.

[24] P. E. Zhidkov. On an invariant measure for a nonlinear Schrödinger equation. *Soviet Math. Dokl.*, 43:431, 1991.

[25] B. Bidegary. Invariant measures for some partial differential equations. *Physica D*, 82:340, 1995.

[26] J. L. Lebowitz, H. A. Rose, and E. R. Speer. Statistical mechanics of a nonlinear Schrödinger equation. *J. Stat. Phys.*, 50:657, 1988.

[27] J. Bourgain. Periodic nonlinear Schrödinger equation and invariant measures. *Commun. Math. Phys.*, 166:1, 1994.

[28] W. A. Strauss. *Nonlinear Wave Equations*. AMS, Providence, RI, 1989.

[29] H. P. McKean. Statistical mechanics of nonlinear wave equations IV. Cubic Schrödinger. *Commun. Math. Phys.*, 168:479, 1995.

[30] M. Plischke and B. Bergersen. *Equilibrium Statistical Physics*. World Scientific, Singapore, 1994.

[31] E. T. Jaynes. Information theory and statistical mechanics. *Phys.Rev.*, 106:620, 1957.

[32] C. Canuto, M. H. Hussaini, A. Quateroni, and T. A. Zang. *Spectral Methods in Fluid Dynamics*. Springer, Berlin-Heidelberg, 1988.

[33] D. Revuz and M. Yor. *Continuous Martingales and Brownian motion*. Springer-Verlag, Berlin-Heidelberg, 1991.

[34] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.

[35] R. Adler. *The Geometry of Random Fields*. Wiley, New York, 1981.

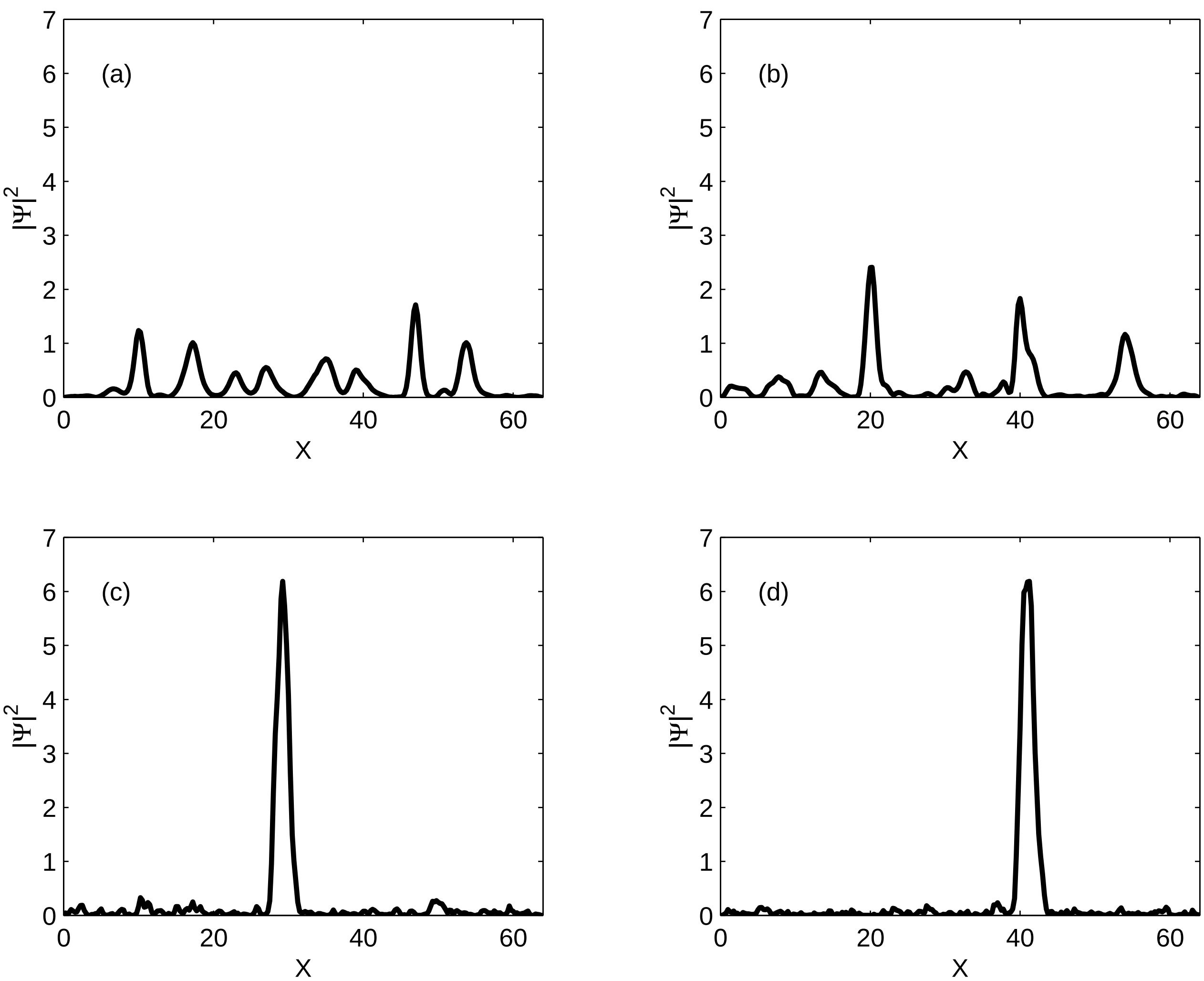
[36] T. Cazenave and P. L. Lions. Orbital stability of standing waves for some nonlinear Schrödinger equations. *Commun. Math. Phys.*, 85:549, 1982.

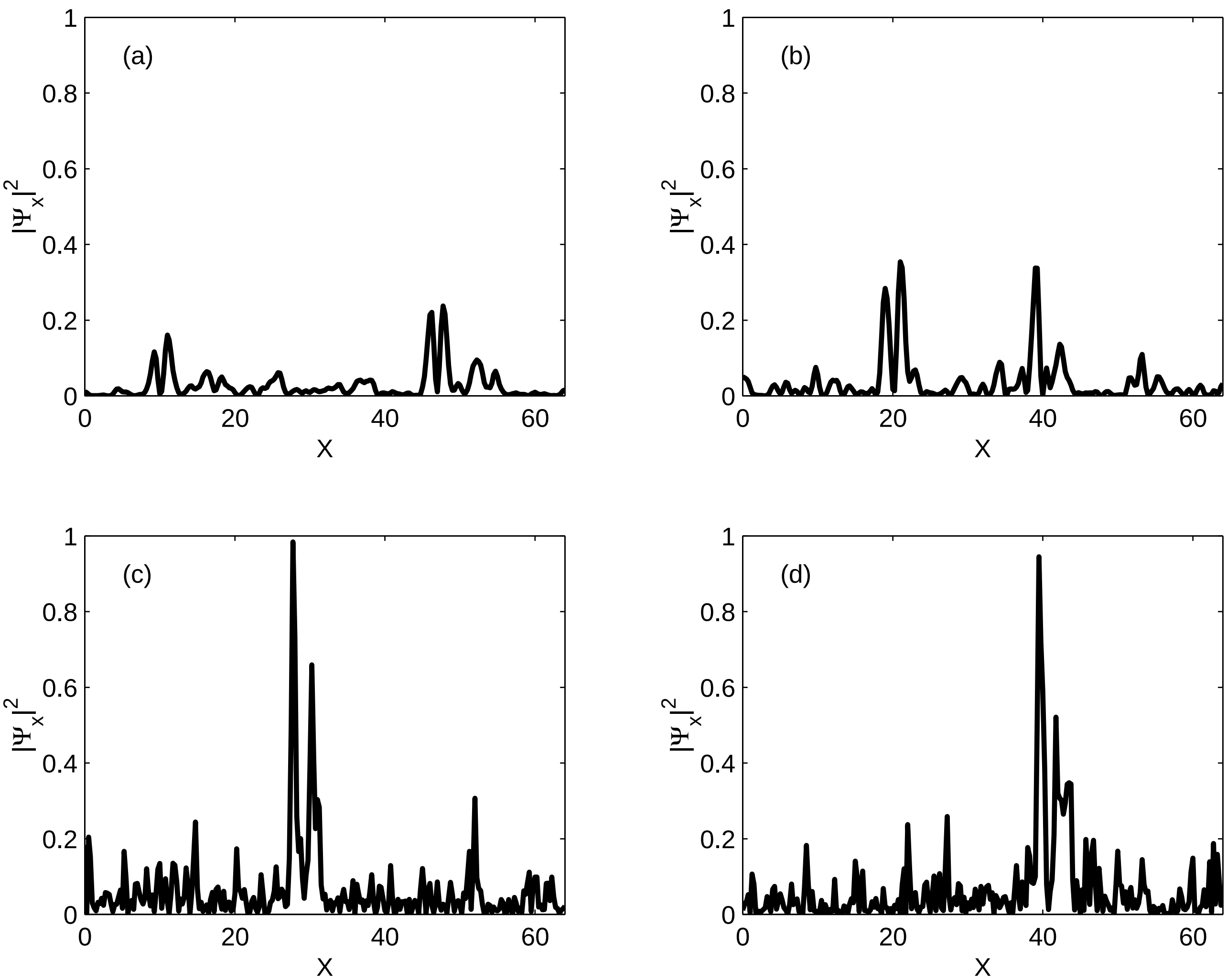
[37] R. Jordan, B. Turkington, and C. Zirbel. On a mean-field statistical theory for a class of NLS equations. *Los Alamos Technical Report*, 1999.

[38] R. Adams. *Sobolev Spaces*. Academic Press, New York, 1975.

Figure 1: Evolution of $|\psi(x,t)|^2$ for the NLS equation $i\psi_t + \psi_{xx} + \psi|\psi|^2/(1+|\psi|^2) = 0$ on the spatial interval $[0, 64]$, with homogeneous Dirichlet boundary conditions imposed. The initial condition is $\psi(x,0) = 0.5 \tanh(x) \tanh(64-x)$ plus a small random perturbation. The integration scheme used is the split–step Fourier method with 512 Fourier modes. The relative deviations $(N(t)-N^0)/N^0$ and $(H(t)-H^0)/H^0$ of the particle number $N(t)$ and the Hamiltonian $H(t)$ from their initial values N^0 and H^0 are less than 3×10^{-8} and 8×10^{-4}, respectively, for the duration of the simulation. (a) t=100; (b) t=1100; (c) t=100100; (d) t=200100.

Figure 2: Evolution of $|\psi_x(x,t)|^2$ for the same equation, boundary conditions, and initial condition as in Figure 1. (a) t=100; (b) t=1100; (c) t=100100; (d) t=200100.


(a)
(b)
(c)
(d)
$|\Psi|^2$
X
0
20
40
60
1
2
3
4
5
6
7



(a)
(b)
(c)
(d)
|Ψx|2
X
0
0.2
0.4
0.6
0.8
1
20
40
60